DOCUMENT RESUME

ED 127 930 IR 003 892

AUTHOR	Abrams, Marshall D.; And Others
TITLE	Measurement of Computer Communication Networks.
INSTITUTION	National Bureau of Standards (DOC), Washington, D.C. Inst. for Computer Sciences and Technology.
SPONS AGENCY	National Science Foundation, Washington, D.C.
REPORT NO	NBS-TN-908
PUB DATE	Jul 76
NOTE	94p.
AVAILABLE FROM	Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 (SD Catalog No. C13.46:908; $1.55)
EDRS PRICE	MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS	Bibliographies; Communications; Computer Programs; *Computers; Costs; Data Processing; Evaluation; Evaluation Criteria; *Evaluation Methods; Information Science; *Information Systems; *Measurement; Measurement Instruments; Measurement Techniques; Networks; On Line Systems; Performance Criteria; Program Evaluation; Research Methodology; *Telecommunication
IDENTIFIERS	Computer Networks; Interactive Computer Systems

ABSTRACT

Measures, tools, and techniques applicable to the performance measurement of computer communication networks are described for technicians who procure computer services from a remote access network. Cost considerations are discussed as a major component of evaluation, and measurement and evaluation methodologies are surveyed. External measurement techniques that encompass both the intermediate user-network interface and the distant interface between user and the serving computer system are recommended. An 80-item bibliography concludes the report. (CH)

****
* Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. Nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original. *
****

ED127930

Measurement of Computer Communication Networks

Marshall D. Abrams

Siegfried Treu

Robert P. Blanc

Computer Systems Engineering Division
Institute for Computer Sciences and Technology
National Bureau of Standards
Washington, D.C. 20234

Sponsored by the
National Science Foundation
1800 G Street, N.W.
Washington, D.C. 20006

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

U.S. DEPARTMENT OF COMMERCE, Elliot L. Richardson, *Secretary*

Dr. Betsy Ancker-Johnson, *Assistant Secretary for Science and Technology*

NATIONAL BUREAU OF STANDARDS, Ernest Ambler, *Acting Director*

Issued July 1976

IR003892

National Bureau of Standards Technical Note 908
Nat. Bur. Stand. (U.S.), Tech. Note 908, 95 pages (July 1976)
CODEN: NBTNAE

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1976

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402
(Order by SD Catalog No. C13.46:908) Price $1.55 (Add 25 percent additional for other than U.S. mailing).

TK 003815

FOREWORD

This report is one of a series of publications produced by the Institute for Computer Sciences and Technology, National Bureau of Standards, under Grant Number DCR 72-01206A05 from the National Science Foundation.

That grant supports a broad program of investigation into the foundations of computer networking in support of scientific and related educational efforts.

A listing of completed and planned publications produced by the Institute under this grant follows:

"Some Technical Considerations for Improved Service to Computer Users," T. N. Pyke, Jr. *COMPCON 73*, Seventh Annual IEEE Computer Society International Conference, February, 1973.

"Computer Networking Technology - A State of the Art Review," T. N. Pyke, Jr. and R. P. Blanc, *COMPUTER*, August 1973.

Annotated Bibliography of the Literature on Resource Sharing Computer Networks, R. P. Blanc, I. W. Cotton, T. N. Pyke, Jr., and S. W. Watkins, NBS Special Publication 384, September 1973.

Review of Network Management Problems and Issues, A. J. Neumann, NBS Technical Note 795, October 1973.

User Procedures Standardization for Network Access, A. J. Neumann, NBS Technical Note 799, October 1973.

Network User Information Support, A. J. Neumann, NBS Technical Note 802, December 1973.

"Primary Issues in User Needs," D. W. Fife, Chapter 10 in *Networks for Research and Education: Sharing of Computer and Information Resources Nationwide*, MIT Press, Cambridge, Mass., 1974.

Computer Networking: Approaches to Quality Service Assurance, R. Stillman, NBS Technical Note 800, January 1974.

Review of Computer Networking Technology, R. P. Blanc, NBS Technical Note 804, January 1974.

Network Management Survey, I. W. Cotton, NBS Technical Note 805, February 1974.

A Guide to Networking Terminology, A. J. Neumann, NBS Technical Note 803, March 1974.

Research Considerations in Computer Networking, D. W. Fife, NBS Technical Note 801, June 1974.

Cost Analysis for Computer Communications, R. P. Blanc, NBS Technical Note 845, September 1974.

"Microeconomics and the Market for Computer Services," I. W. Cotton, *Computing Surveys*, June 1975.

Software Testing for Network Services, R. Stillman and B. Leong-Hong, NBS Technical Note 874, July 1975.

Measurement of Computer Communications Networks, M. D. Abrams, S. Treu, and R. P. Blanc, NBS Technical Note 908.

"A Methodology for Interactive Computer Service Measurement," M. D. Abrams and S. Treu, to be published.

"Network Access Techniques: A Review," R. Rosenthal, *Proc. 1976 National Computer Conference*.

A Review of Network Access Techniques With A Case Study: The NAM, R. Rosenthal, NBS Technical Note (in press).

Software Tools: A Building Block Approach, I. T. Hardy, B. Leong-Hong, D. W. Fife, NBS Technical Note (in press).

Contents

		Page
1.	Introduction	1
2.	Performance Measurement	4
2.1	Measures	4
2.1.1	Measures of Time	5
2.1.2	Measures of Length	10
2.1.3	Multiplicities and Frequencies	13
2.1.4	Measures of Rate	14
2.1.5	Ratios and Other Measures	15
2.2	Tools and Techniques	17
2.2.1	User-System Interaction	17
2.2.1.1	Interaction Models	17
2.2.1.2	Measurement Devices	18
2.2.1.3	Data Processing	19
2.2.2	Measurement Under Controlled Conditions	20
2.2.3	Sample Experimental Data	21
2.3	Approach to Usability	23
2.3.1	Action Primitives	28
2.3.2	Transformation to Command Language	29
2.3.3	Associativity in Mental Work	31
2.3.4	Appropriate Measurements	32
3.	Cost Considerations	33
3.1	Costing Alternatives	34
3.1.1	Traditional Factors	34
3.1.2	User-Oriented Factors	35
3.1.3	Hybrid Approach	36
3.2	Communications Costs	37
3.2.1	Typical Cost Factors	38
3.2.2	Actual Line Utilization	39
3.2.3	Data Collection Techniques	41
3.2.4	Data Traffic Characterization	44
4.	Summary and Conclusions	45
Bibliography		48
Appendix A		55
Appendix B		57
Appendix C		60
Appendix D		66
Appendix E		69
Appendix F		74

MEASUREMENT OF COMPUTER COMMUNICATION NETWORKS

Marshall D. Abrams
Siegfried Treu
Robert P. Blanc

Abstract

This report is concerned with aiding those responsible for the procurement of computer services from a Remote Access Network by providing a description of the measures, tools, and techniques applicable to the performance measurement of computer communication networks. Cost considerations are discussed as a major component of evaluation. Measurement and evaluation methodology are surveyed, including various operational tools and techniques. Some exemplary data are also presented. Although the constituents are already present, a neatly packaged methodological product, perhaps in the form of a well-structured user's guide to network performance measurement, is not yet available.

Key words: Communications networks; computer networks; cost; interactive service; measurement; measures; performance; usability.

1. INTRODUCTION

This report identifies key criteria to be used in specifying, for selection and service quality assurance purposes, the performance of a computer communications network, and identifies methodologies for making quantitative measurements.

The report has three main objectives:

1. To identify key criteria and associated measures to be used in specifying the performance of a computer communication network or network service for initial selection purposes.

2. To develop a set of measures for monitoring the performance of procured network communications service.

3. To define a methodology for objectively comparing, on a performance and cost basis, competing computer communications network services.

The criteria and measures developed should be useful to the users, and potential users, of large national computer communications networks. If the user wants one or more computer services to be made available to many terminal locations with national distribution, a number of options for satisfying his computer communications needs present themselves. Assuming he is in an appropriate position and can afford it, he can design and implement his own computer network or network-based service. On the other hand, an existing network may be available for purchase. Furthermore, he could contract for a customized version of an existing computer network or arrange for the connection of his local computer terminal to such an existing facility. Various other alternatives or shades of dependency on, or modes of access to, computer communications networks are possible.

Many of the criteria, which may be thought of as dimensions or vectors, developed in this report will be applicable to the option of connecting to an existing network, and will be especially useful in carrying out objective measurements of the effectiveness of alternative computer communications networks for the user's needs. Only minimal attention is given to the intricacies of actually designing and implementing a computer communications network.

The network user usually does not have to be concerned with, or to even be aware of, the intricacies of communications hardware/software and is not really assisted by the types of "internal" performance measures which have been widely developed. A user would benefit more from a methodology for making quantitative external measurements considering both performance and cost factors to assist in the objective comparison of competing services and viable alternatives. Such is the emphasis of this report.

The contents of this report are equally applicable to any existing communications network where it is desired to measure performance with respect to the satisfaction of user needs. The emphasis is on measurement for the procurement of computer communication services and the monitoring of those services thereafter; the latter is applicable to all computer communications networks.

In the procurement of network services, it has been difficult to determine and clearly specify meaningful and useful performance criteria. The work described in this report is to improve that situation by generating performance-oriented specifications for purposes of application to objective measurement and subsequent evaluation of computer communications network services.

A particular orientation or philosophy is maintained in this effort. Performance measurement techniques that indicate percent utilization of a CPU or some other internal measure of performance are comparatively meaningless to a remote user (or user population). The user is more concerned with the visible amount of work performed for him per unit time and the cost of the computer communications facility for the entirety and for various subsets of that work. Included among the costs of the facility are those attributable to computer hardware, software, and transmission facilities, both within the communications network itself, and special adaptations for host computers which are necessary to allow them to communicate through the network. Examining performance from a user's viewpoint leads to external measures of performance and to new measurement techniques. Techniques currently being investigated include recording the dialogue between a remote terminal user and a host computer system on a network through the use of a new tool, the "Network Measurement System."

The application of this machine combined with a set of analysis packages can lead to a methodology for objectively comparing and analyzing, on a quantitative basis, the services offered by networks. Similarly a methodology can be developed for the subsequent monitoring of selected networks. Section 2 of this report concentrates on the discussion of these techniques and methodologies.

Given useful measures of performance, attention can then be given to cost/performance considerations. It is possible to identify several categories of cost factors. Section 3 covers both the traditional variety and others of particular pertinence to this report, including costs of using communications facilities for connecting a terminal to a network and costs for interconnecting many terminals to host computers through use of a network. The communications cost factors have been partially identified for existing and proposed computer communications networks, including such factors as the cost of installing and operating the host computer interfaces and terminal interfaces, and the costs of transmitting traffic through the network. In Section 3, methods are discussed for using these cost factors, consistent with new techniques for the controlled measurement of user traffic demands, to derive cost estimates for networking for specific configurations.

Another specialized area related to measurement of the cost/performance of a network is the measurement of the usability of a network. Through a network many different services may be available, with different access or command procedures for performing a wide variety of functions. Determining the effectiveness of particular access procedures through quantitative measures is an area in need of attention. In Subsection 2.3 a methodology is proposed, for comparing the usability of alternative access procedures using a technique for quantitatively measuring user "think" time in relation to alternative commands.

In combination, the proposed tools and methodologies may lead to a detailed plan for measuring competing computer communications facilities to determine capabilities, performance, and costs to aid in the selection of network services. Unfortunately, the present state-of-the-art does not permit the combination of these components into prescriptions. Section 4 summarizes the material of the previous sections, and through the identification of the key points, arrives at the conclusions and recommendations possible at this time.

2. PERFORMANCE MEASUREMENT

As indicated in the Introduction, the emphasis of this report is on providing the prospective user or procurer with some methodology for making quantitative external measurements for purposes of objective comparison of computing services. This section addresses that methodology in terms of appropriate network performance measures and measurement tools and techniques.

2.1 Measures

In trying to develop a suitable user-oriented methodology for measuring network performance, one can argue that the real demands that are made by users of available computer networks and which are capturable by means of the tools and techniques to be described later represent the best objective approximation to current user "needs." This argument displays the strengths of avoiding user opinions on their needs, along with various related value judgments, and of relying principally on analysis of objective data about user-network interaction.

Given that rationale, it is necessary to view the totality of collectible data representative of user demands and needs and to elicit, define, or prescribe those measures which are somehow representative of interactive network performance characteristics. To be as exhaustive as reasonably possible, one must conceptualize (based preferably on experience) the tasks facing various user types and those network features which they might consider as important, either directly or, at least, indirectly.

This section, which includes several tabulations of measures, resulted in part from such deliberate consumer-oriented thought processes. It was also based on other work reported in the literature, to be cited as appropriate.

2.1.1 Measures of Time

By the very nature of on-line, real-time user-computer interaction, the variety of timing data pertaining to that interaction is very conducive to measurement and analysis. The question to be answered is: what time-based system characteristics are either individually or collectively observable by the user while carrying implications on some aspect of network performance? Whatever they are, it may be possible to define applicable measures, or formulas of pertinent timing variables, the values of which can be meaningfully interpreted as good, bad, or otherwise.

Table 1 presents a list of such measures. Its first entry, namely system delay, represents one version of what has been frequently and variously considered as response time. A problem with this measure is disagreement on precise definition of the measured time interval. System delay time (later to be called stimulus-acknowledgement delay) is taken to be the time from the user's carriage return to the time of arrival of the first character (maybe control character) from the system. Several other valid alternatives have been suggested and studied, and these will be given special, separate attention in the next subsection.

The importance of response time to interactive service is generally acknowledged. Dexter [1973] has gone as far as to state that "Response time is a surrogate for value because, in selecting among time-shared services, the user in general will pay more for faster response times...." Some investigators [e.g., Carbonell et al., 1968] have discussed response time by drawing analogies between user-computer communication and human-human conversations. After all, just as a human conversation becomes non-existent when gaps between talk spurts become too long, an interactive service

Table 1. Time-Based Measures

NUMBER/NAME		VARIABLE/FORMULA		UNIT OF MEASURE	MEANS OF MEASUREMENT	COMMENTS
		Single	Cumulative			
	Time-Based					
1.	System Delay	t_{1i}	$\sum_i^n t_{1i}$	sec	1, 2	Where $1 \le i \le n$, for n system responses per session
2.	System Transmit	t_{2i}	$\sum_i^n t_{2i}$	"	"	Also called printout or output display time
3.	Acknowledgement Delay	t_{3i}	$\sum_i^n t_{3i}$	"	"	Time from input carriage return to first system reaction
4.	Acknowledgement Transmit	t_{4i}	$\sum_i^n t_{4i}$	"	"	Reflects length of that reaction (which may be entire system response)
5.	User Delay	t_{5i}	$\sum_i t_{5i}$	"	"	Includes user read, think, etc. time
6.	User Transmit	t_{6i}	$\sum_i^n t_{6i}$	"	"	Includes time required for user input typing
7.	User Task	t_{7j}	$\sum_i^n (t_{1i}+t_{2i}+..+t_{6i})$	"	1, 2, 3(or 4)	Where $1 \le n_1 \le i \le n_2 \le n$, and the jth task is delimited by transactions n_1 through n_2
8.	Selected Subsession	t_{8k}	"	"	"	Where i is restricted to those contiguous or non-contiguous transactions defining the kth special subset
9.	Interprocess Transfer	t_{9p}	"	"	"	Where i is restricted to those transactions involving movement between different software or hardware processors
10	Session Start, Finish	t_{10}, t_{11}		"	1, 5	
11.	Total Session	$t_{11}-t_{10}$	$\sum_i^n (t_{1i}+t_{2i}+..+t_{6i})$	"	1, 2, 5	Which may consist of one or more successive user tasks, as well as use of one or more processors

MEASUREMENT OR IDENTIFICATION MEANS:

1. Application of Network Measurement Machine (see Section 4)
2. Software analysis of collected data
3. Software search and identification within alphanumeric data base
4. Human search and identification within alphanumeric data base
5. Access to available accounting records

becomes noninteractive if the computer fails to respond quickly enough. But, then, it may be that attention to response time in itself has been excessive. Kamerman [1969] has criticized its use "as the prime measure of time-sharing-system performance." He feels that it is weak in leaving certain questions, e.g., what is being responded to, unanswered. He does acknowledge, however, that his "conversational throughput" measure subsumes response time.

Most of the other measures listed in Table 1 have been comparatively neglected. The system transmit or printout can be usefully viewed as one indication of system "verbosity," particularly when restricted to a certain class of responses. Perhaps the user is being inundated by much unnecessary or redundant verbiage.

The need for system acknowledgement within tolerable delays has been studied [Miller, 1968] and is widely recognized. Some difficulties with its definition and measurement will be covered in Subsection 2.2.1. But little has been done to interpret the meanings of the user-contributed time components in an interactive session. Subsection 2.3 will pay some separate attention to studying these components.

Total session or console time, being easily obtainable, has been logged by many people and for various purposes. Besides being useful for calculating certain rates and ratios (in Subsections 2.1.4 and 2.1.5), it does give a general image of system availability to a user, especially if averaged over a number of sessions. However, it says little if anything about what and how much the user may in fact be accomplishing during a session.

With that in mind, it is helpful to subdivide a session into tasks, subsessions, interprocess transfers, and perhaps other similar entities or points of closure. These can then give rise to measures of service rates (e.g., in terms of tasks completed), or "single-user throughput," as well as measures of nonproductive, intermediate transfer points.

It should be recognized that all of the measures included in Table 1, as well as Tables 2 and 3, are purposely intended to apply to one user conducting one interactive session with a system via a network. That is to say, assuming that the total system load can be established and maintained at some reasonably stable level, the listed measures should be applicable to a single user in attempting to access computing services, regardless of what the remaining "anonymous" user population does and thinks.

Table 2

LENGTHS AND MULTIPLICITIES

NUMBER/NAME	VARIABLE/FORMULA		UNIT OF MEASURE	MEANS OF MEASUREMENT	COMMENTS
	Single	Cumulative			
Length-Based					
12. System Response	L_{1i}	$\sum_{i}^{n} L_{1i}$	# char	1,2 (or 5)	This gives an indication of system verbosity
13. User Message	L_{2i}	$\sum_{i}^{n} L_{2i}$	"	"	This can be used in assessing perhaps excessive command structures
14. Acknowledgement	L_{3i}	$\sum_{i}^{n} L_{3i}$	"	"	
15. Session Length	L_4		# transactions	"	One transaction = one message + acknowledgement + response
16. Task Length	L_{5j}	$\sum_{j} L_{5j}$	"	1,2,3 (or 4)	Where the task, e.g., with an editor or a compiler, must be defined
17. Subsession Length	I_{6k}	$\sum_{k} L_{6k}$	"	"	This can be used in studying processor/command utilization
Multiplicity					
18. Command Use	m_{1r}	$\sum_{r}^{s} m_{1r}$	#	1,2,3 (or 4)	The number of times a specific (the rth) command (type) is used
19. Task Succession	m_2		# tasks	"	Number of different tasks carried out in one session
20. Processor Variety	m_3		# processors	"	Number of different processors used in a session
21. Processor Call	m_{41}		#	"	Number of times a specific processor is called
22. Processor Switch	m_5	$\sum_{1}^{m_3} m_{41}$	#	"	Number of times the user switches among the m_3 processors
23. System, User Error	m_6, m_7		# errors	"	Distinguishing a system- from a user-caused error
24. Faulty Transaction	m_8		# transactions	"	Number of transactions with one or more errors, regardless of fault

Table 3

RATES AND OTHER MEASURES

NUMBER/NAME	VARIABLE/FORMULA	UNIT OF MEASURE	MEANS OF MEASUREMENT	COMMENTS
Rate-Based				
25. Character Arrival	$r_1 = \Sigma L_{1i}/\Sigma t_{2i}$	char/sec	1,2	This example rate is speed over entire session
26. Character Departure	$r_2 = \Sigma L_{2i}/\Sigma t_{6i}$	"	"	Can also be viewed as a global typing rate
27. Interaction	$r_3 = L_4/(t_{11} - t_{10})$	trans/hr	"	This gives a frequency of user-system interaction
28. Task Completion	$r_4 = m_2/(t_{11} - t_{10})$	tasks/hr	1,2,3 (or 4)	This is interpretable as an individual user task throughput measure
29. Processor Interchange	$r_5 = m_5/(t_{11} - t_{10})$	switch/hr	"	
30. Total Error	$r_6 = \frac{m_6 + m_7}{t_{11} - t_{10}}$	error/hr	"	
31.-39. A number of other definable rates				
Ratios				
40. Relative Response	$s_i = L_{1i}/L^*_{1i}$	-	1,2	Where L^*_{1i} is obtained under lower (maybe optimum) system load
41.-49. Other similarily relative measures				
50. Reliability	$s_2 = (1 - m_8/L_4)$	-	1,2,3 (or 4)	The fraction of undisturbed transactions for an entire session
51. Command Utilization	$s_3 = m_{1r}/L_4$	-	"	Compares the r^{th} command type against total number of commands
52. User Idleness	$s_4 = \frac{\Sigma(t_{1i} + t_{3i})}{t_{11} - t_{10}}$	-	1,2	Reflects the fraction of time the interactive user is waiting

53. etc. The well-known Mean, Median, Std. Deviation, etc., measures applicable to many of the entries in these tables.

This single-user, single-session orientation, which is born out by both single and cumulative (where appropriate) variables and formulas listed in the tables, is also characterized in Figure 1. Such orientation is clearly suitable for the application of measurement tools to individual user-system interactions. It is not to imply that measurement and evaluation should stop with the single user. There is obviously purpose and strength, as well as potential for more meaningful results, in studying user (sub)groups collectively, particularly for the types of performance monitoring and procurement purposes of interest in this report. Figure 1 suggests two higher levels of measurement orientation.

As is true for single-user applicability, with many results analyzable per session as well as across a number of sessions, the collective user approach can rely on the well-known statistical indicators, including the mean, median, standard deviation, 90 percentile, and others. These are not explicitly identified in the tables as to their appropriateness. It was deemed sufficient to list only the fundamental measures which represent the basic constituents for calculating the standard statistical values.

As indicated earlier, it is generally agreed that some measure of system responsiveness is fundamental. But there is no consensus to the definition of response time. Since this time-based measurement is basic and intuitively appealing, a brief discussion of the various definitions found in the literature is given in Appendix B. Each one represents an author's conclusion of how to measure a computer's response to a service request stimulus by the human user.

2.1.2 Measures of Length

In view of the essentially serial nature of (present-day) man-computer communications, one might expect a high correlation between the time it takes to transmit a message (or user-initiated stimulus) or system response and the corresponding message or response length. That is of course not generally true for human-typed messages, except perhaps when very short commands are involved [Holdsworth et al., 1973], and even the system which can transmit at a steady, consistent rate may have sporadicity and other problems.

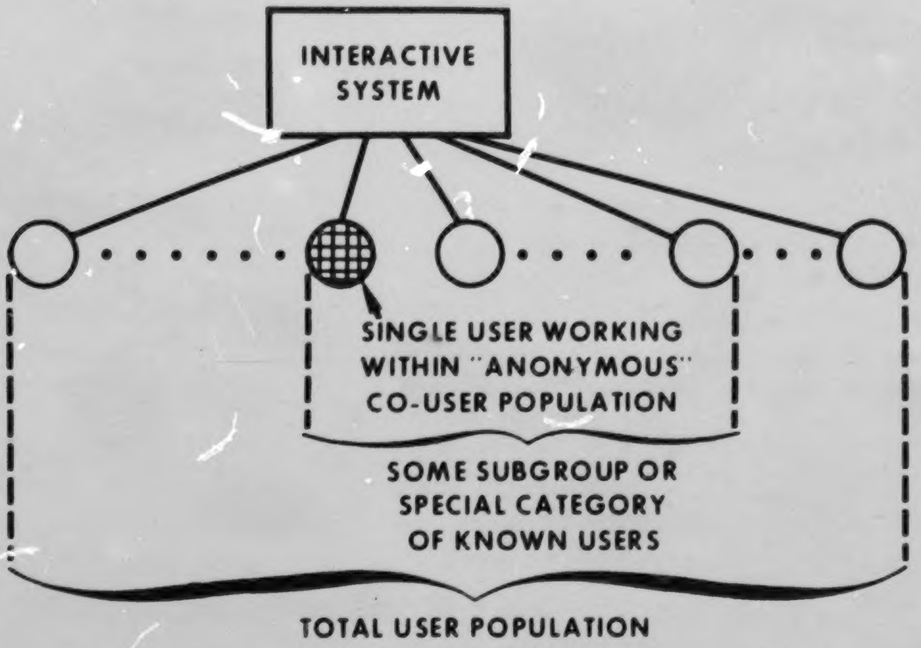

INTERACTIVE SYSTEM
SINGLE USER WORKING WITHIN "ANONYMOUS" CO-USER POPULATION
SOME SUBGROUP OR SPECIAL CATEGORY OF KNOWN USERS
TOTAL USER POPULATION

Figure 1. Single and Collective Applicability of Measures

Thus, it is worthwhile to consider pure length as a service parameter of interest. In fact, it may very well be that, aside from any related timing considerations, the lengths of system response and required user message possibly indicate unnecessary system verbosity and undesirably complex command structures or syntax. Although such conclusions have been difficult to ascertain, except by means of direct subjective comparison with other similar systems which are more or less verbose and complex, the approach described in Section 2.3 has the potential for enabling or facilitating them.

Table 2 also displays some length-based measures expressed in terms of numbers of transactions. One transaction is defined to be one stimulus-acknowledgement-response group. The total session or conversation then consists of a certain number of transactions. More interestingly though, with the view of trying to elicit some characteristic(s) implying "work accomplished" or throughput, it should be possible to identify either a contiguous sequence or an interspersed, noncontiguous sequence of transactions as representative of a user task. The latter can also be called by the generic term "subsession."

How are such tasks identified? In the legend of Table 1, five different means of measurement are outlined. The first two means are self-explanatory. The third is subject to feasibility constraints involving the possible need for exhaustive, complex searching of the user-system dialogue text. To carry out software- based searching for purposes of task or subsession identification and corresponding data compilation, an effective and reasonably compact set of search criteria must be available. If the system is too inconsistent in its command structures and in distinguishing features among its different software processors, and if, in addition, considerable interspersement of tasks is possible (e.g., by direct accesses to the editor), then a human assistant may be essential. A knowledgable person could review the user-system dialogue, perhaps by scrolling it across a display screen, and simply mark the delimiting task boundaries or constituent transactions according to some established criteria. The resulting marked data base can then be processed further under software control.

The fifth means for identifying or collecting data, as listed in Table 1, appears to straddle the boundary between external and internal measurement. If the system itself maintains accurate records of such measures as total session time, which can be interrogated interactively by the user (or a user program), then these represent an alternate or verifying source of data. Of course, system-controlled accounting is normally most concerned with utilization factors for CPU, memory, input/output and other equipment. This category of accounting data, if interrogatable, can also serve a useful but different purpose by providing information about time-dependent variation in computer use. The ability to ascertain system and network load with respect to time-of-day and other factors is an essential prerequisite to achieving meaningful measurement results.

2.1.3 Multiplicities and Frequencies

A number of service-related characteristics lend themselves to simple tallying or counting to arrive at multiplicities or frequencies of occurrence during a user-system interaction. The only problem may be the identification of what is being counted. A human identifier may again be required, as discussed previously.

Among the multiplicity measures listed in Table 2, the most troublesome are probably those involving system or user errors and fair assignment of fault. The others, however, should be relatively easily obtainable.

What do these measures have to do with user-observable network-based services? "Command use" carries implications for the nature and quality of the available command repertoire. If a particular command (set) is used very frequently, it naturally leads to comparison against infrequently or unused commands (see also "command utilization" ratio in Table 3). Perhaps the service can be improved by greater emphasis on and more efficient handling of frequently used commands. Maybe a "flexibility" in command procedure is called for, enabling the users to write their own, tailored commands [Doherty et al., 1972].

"Task succession," which is necessarily dependent on lengths and types of constituent tasks, implies something about how well the system is equipped and functioning for fostering the user inclination (or persistence) to get several tasks done in one session, as opposed to becoming frustrated and quitting sooner. The measures involving processor variety, calling, and switching may in fact play a related role. They generally convey a "selectability" and "mobility" among different processors in a system which can have a considerable bearing on how much a user can accomplish.

2.1.4 Measures of Rate

Several of the fundamental measures involving a unit of length or simple multiplicities gain added significance when related to time. Thus, the total number of characters transmitted by the system can be divided by their collective transmission time to result in a measure of character arrival rate (see Table 3). It is easily seen that this could be developed into a measure of system "character arrival sporadicity" by comparing actual character arrival against the standard character transmission rate and perhaps weighting it appropriately. Results could have important implications on the network communications facility.

A similar measure is available in "character departure" rate. Aside from behavioral considerations involving user typing rate, an important service-oriented question is whether different character departure rates are treated differently by the communications network and/or computer system, and, if so, what can be done about it.

The interaction rate is clearly a measure which reflects "conversational ability." If the number of transactions is comparatively low, the user might legitimately view the service to be only semi-interactive. Several investigators have dealt with interaction rates, although in different ways. Streeter [1972], who considers user productivity to be proportional to the number of interactions per unit time, has defined the following:

Interactions per hour = 3600/(S+U)

where S = mean system response time (in sec)

U = mean user response time (in sec)

Lassettre and Scherr [1972] have a similar-looking measure but with somewhat different interpretation:

Total interactions processed = n/(U+R)

where n = average number of actively interacting time-sharing users

U = average user time

R = average response time

The latter version is clearly oriented to an entire user group (see Figure 1) as opposed to only the individual user.

One other rate-based measure warrants special mention. The "task completion" rate, based on the task succession count of Table 2, can be interpreted as a minimal measure of single-user throughput. It is minimal in the sense that it reflects only one user's accomplishments without taking other users and resulting system load into account. Assuming that the task identification requirement is workable, the task completion measure could easily be developed into a collective measure representing any selected user group. In either case, it would be consistent with the conversational "throughput" measure defined by Kamerman [1969], namely "performance of a given amount of work in a given amount of time."

2.1.5 Ratios and Other Measures

A number of ratios have been suggested in the literature as useful measures. One is called the "interference ratio of response time," defined to be actual response time divided by optimal response time [Dexter, 1973]. This relative response measure is indicative of the problems existing in measuring and evaluating any time-sharing service: you must have a reliable basis for characterizing pertinent system load. Otherwise, the results of repeated measurements may be quite different under different sets of circumstances but may be inexplicable for want of understanding the effects of system loading.

Many "relative" measures are available. Some, as the one defined above, represent attempts to characterize different conditions within the same system. Practically all of the measures listed in Tables 1, 2, and 3 can be used for comparison against other network-based systems.

The tabulated measures also give rise to some special combinations. Reliability, for example, when viewed within a single session by a user, can be interpreted as the fraction of undisturbed transactions (which could possibly be weighted according to lengths of any disturbances). Another interesting measure is that of "user idleness." It relates the time expended by the user waiting for system acknowledgements and system responses to total session time. It can be argued that the (perhaps highly-paid) user who is forced to waste too much time in a pure waiting mode is not being adequately serviced, at least from the user management's standpoint.

Still other measures do not fit into the "single-user, single-session" orientation which is predominant in this paper. They are nevertheless useful. The characteristic of system "accessibility," for example, is amenable to a "pre-session" measure of the user's ability to get in and utilize the system services [Grubb & Cotton, 1975]. A simple ratio of successful to attempted accesses, over a period of time, is a possible candidate. Carbonell et al. [1968] have addressed the accessibility problem of getting a busy signal and suggested two factors as possible determinants of the caller's degree of frustration with the system: (1) percentage of time a busy signal occurs, (2) expected length of waiting time before getting access.

The latter factor naturally precipitates more general questions about system service "predictability." It would be very helpful if the service consumer were provided a reliable probability of gaining access and, thereafter, of experiencing service above some specified level. Much useful work on such consumer aids is left to be done. An interesting related example, however with orientation to internal system performance, is the "probability of no requests for service in the system," given n on-line users, as suggested by Lassettre and Scherr [1972]:

$$P_0 = 1/\left(\sum_{i=0}^{n} (n!/(n-i)!)(S/U)^i \right)$$

where S = average service time per instruction

U = average user time

Finally, it is, in fact, possible to talk about and measure environmental effects such as embodied in various management-imposed policies and procedures, on the network-based computer service. Streeter [1972] is an example investigator in this area. Without reviewing the details of his reported work, he compared some alternative policies on modes of system use and in essence concluded that policies should encourage the use of interactive systems such that user benefits exceed the sum of all cost factors by at least some management-established minimum. Among the resulting advantages of "policies that encourage the right choice of computer service" are more effective use of equipment and, more importantly, "helping people become more effective and efficient."

2.2 Tools and Techniques

Probably the most difficult constituent to contend with and control in measurements of the user-system interface is the human user. The inconsistencies and idiosyncrasies of user behavior are well known but troublesome to predict or even explain. For this and other reasons it is desirable to stratify the measurement methodology to encompass both options of real and simulated user participation.

Accordingly, we have dichotomized the description of measurement tools and techniques. The next section presents those which are in fact applicable to the operational situation in which real human users utilize the services of an interactive, network-based system. Special additional tools are required to enable similar measurements relative to simulated user demands. These are discussed in Section 2.2[illegible] with further elaboration on the pertinent experimental control considerations.

2.2.1 User-System Interaction

When actually performing measurement and evaluation, it is not only necessary to understand the object of measurement, but also the level at which measurement is taking place. For this purpose, it is useful to present the interactive process as a conceptual model. Several such models for the man-computer interaction should be considered.

2.2.1.1 Interaction Models

As discussed in Abrams [1974], the simplest model is a stimulus-response couplet wherein the human computer user issues the stimulus and the computer issues the response. Except for the acknowledgement, this model is sufficient for the interaction with a half duplex computer system which does not queue successive service request stimuli (e.g., IBM System/360 TSO). In this discipline, the human user is forced to alternate his stimuli with the computer's responses; violations of this discipline are simply ignored (when a non-locked keyboard makes them possible at all). Some half duplex systems provide for command queuing (e.g., Univac 1108 Exec 8), thereby introducing the possibility of non-contiguous interleaving of stimuli and responses. The modeling of this situation is more complex only in the association of a response with its stimulus [Abrams et al., 1973]. Further complexity in modeling the man-computer interaction is introduced by highly interactive full duplex systems which provide command and sub-command prompting (e.g., DEC System-10 TENEX). In this situation, the model must accommodate stimuli, sub-stimuli, responses, and sub-responses as well as the proper associations among them.

The simple stimulus-response model has proved to be unsatisfactory for those computer systems which issue a line feed (LF) whenever the user types a carriage return (CR). This action is common when ASCII [1968] terminals are employed for communication. Treating the LF as the beginning of the computer's response leads to the anomalous interpretation that a long response time is in fact a short one. A simple extension is the incorporation of three states in the model, which is renamed the stimulus acknowledgement-response (SAR) model. The new state, the acknowledgement, includes the LF and other non-printing characters which accompany it. One alternative definition is that the acknowledgement includes all characters following the stimulus which do not convey any information to the user.

2.2.1.2 Measurement Devices

By now it is fairly common to instrument a computer system for purposes of refinement, debugging, and configuration analysis. Common techniques, employed individually or in combination, include hardware modification of the computer, attachment of external hardware monitors without modification to the main frame, and introduction of monitoring software in the executive or as user programs. For further information see Miller [1972].

Exception notwithstanding, the data collected by the mentioned techniques are not generally suited to analysis of service rendered. While exceedingly appropriate for measurement of internal performance, the state variables collected are too microscopic or introverted to be applicable to service measurements. It is, of course, quite interesting to correlate such internal measurements with the external service measurements. As discussed by Saltzer and Gintell [1970], intuition frequently fails in complex computer performance analysis; the measured interrelation of internal and external performance should be quite useful in hardware configuration balancing as well as software algorithm enhancement. Furthermore, many of the internal performance measurement systems are specific to a particular model of computer hardware and/or operating system. In many cases they are implemented internally to the system, thus perturbing the object of measurement; many are also regarded as proprietary tools. This report refers primarily to measurement tools and techniques in the public domain, or at least described in the literature, applicable to a broad range of computer equipment with the understanding that similar methodology may exist elsewhere.

A prototypical tool for service measurements exists in the Network Measurement Machine (NMM) [Rosenthal, Rippy and Wood, 1976]; similar features are implemented in other systems such as the Remote-Computer-Controlled Hardware Monitor (RCHM) [Morgan et al., 1974]. Based on a minicomputer, the NMM is equipped with interfaces, a programmable high precision clock and a recording medium which enable it to identify, time stamp, and record communications traffic between the human user and the serving computer system. The identification referred to above is only concerned with the source (user or computer) of the character.

The NMM may be used to gather information solely relevant to the consumer of computer services. In this mode, the NMM is connected to the user's terminal. Two methods are described in Appendix C.

2.2.1.3 Data Processing

Our operational procedure has been, firstly, to hypothesize a model (see subsection 2.2.1.1) for describing the physical reality to be measured. The model is an abstraction of a very complicated process involving a human being and highly sophisticated electro-technology. Secondly, some data have been gathered. Then we have tried to take the data and fit it to the model. So long as there are no inconsistencies between the data model and the physical system and we are satisfied with the ability of the model to meaningfully interpret the data, we consider the model to be acceptable. When such inconsistencies are observed it is time for refinement of the model and the repetition of the process. Sometimes more than one model is employed to emphasize different aspects of the measurement. The advantage of multiple conceptual models is the suppression of unwanted detail. Suppression of this detail does not imply that it is irrelevant or uninteresting; but permits concentration on particular aspects individually. As insight is gained it is often possible to combine simple models into a more complicated one with various special cases being possible. The data processing is described briefly in Appendix D and in more detail in Watkins and Abrams [1976].

/

2.2.2 Measurement Under Controlled Conditions

Up to this point, we have defined and described a set of metrics by which to judge the service delivered from an interactive, network-based computer system and the tools, techniques and mechanisms which may be employed in their application. Now we must address the conditions under which measurement is to be conducted.

There are two fundamentally different sets of conditions for measurements. While they are complementary in the sense that they can corroborate or confirm each other, they are sufficiently different in concept that a clear understanding and separation is necessary. The first set of conditions relates to the normal system operation, with the regular distribution of users, user-imposed workload, and system hardware resources. This means that actual workload due to any identifiable user will fluctuate according to whether that user chooses to compute and according to what work he chooses to accomplish. Both random and periodic components characterize this workload; in fact, it may be one object of measurement to determine the extent and nature of these components. For any set of measurements, the hardware and software resources of the server are assumed to remain constant. Measurements before and after a resource change may be employed to determine the effect of that change on service rendered.

When dealing with the normal operating computer system serving real users, we are performing measurement under very complex conditions in which the users are substantially "uncontrollable." Statistically valid samples must be taken; the results are expressed in terms of mean, median, standard deviation, skew, etc. Care must be taken in the design of the data gathering to avoid biasing the data through erroneous sampling techniques.

The second set of conditions is that in which the pertinent variables normally reflecting human users are under the control of the experimenter. In this case, we regard the server computer as an operational or experimental device on which measurements are to be made. Much as in a physics laboratory experiment, we attempt to hold certain independent variables constant, others are changed in a controlled manner, and the values of both the independent and dependent variables are recorded. In this situation it is assumed that if the experiment is repeated, the same results will be obtained. Of course, in this type of experiment there are systematic errors so that even the most perfect repeatably achievable measurement would yield disparate results. Therefore, as before, statistical techniques must also be applied. However, the effects of such factors as individual differences among human users are avoided.

An attempt to perform repeatable experiments to measure deliverable computer service requires that the workload presented to the server be exactly repeatable and controllable. Inconsistency is such that a single human user is incapable of producing such a precise repetition of workload; for a group of users it is even more inconceivable! Furthermore, the expense and logistics of performing experiments with live users may be prohibitive. Therefore, some mechanism is required to produce the conditions of measurement. The obvious mechanism for simulating a group of computer users is another computer. Appendix E explores the conditions which must be met along with several implementations of the class of mechanisms called "measurement drivers." With card-oriented batch processing systems, the repeatable, controlled utilization mentioned above was created by using sets of benchmark programs designed to represent the normal or projected user load on the system (cf. [Joslin, 1965], and [Arbuckle, 1966]).

A word must be said about measurement under mixed conditions, namely when there is some uncontrolled user workload and some controlled driver workload. Measurement under mixed conditions can be quite useful when it is impossible or prohibitive to control the total environment, but when it is desired to measure the service delivered in response to a specific set of workload demands. Inter-system comparison is a prototypical example. In this mixed environment, it is necessary to satisfy both the conditions of statistically valid sampling techniques and the conditions of controlled repeatability in order to obtain meaningful data.

2.2.3 Sample Experimental Data

A substantial methodology for measurement has been described and prescribed in this paper. The value of such efforts has definite limits if the methodology cannot be validated with real measurement work and data analysis.

As implied by the description of measurement tools, the Institute for Computer Sciences and Technology is engaged in a considerable effort in this area. A number of measurement activities, involving the use of the Network Measurement Machine, are currently in progress. These are categorizable as follows:

a. Computer system performance - with respect to the standard internal resources (e.g., CPU, memory, I/O)

b. Interactive system services - the primary focus of this report

c. Communications facilities - to study communication line utilization and time delay characteristics

d. User behavior and characterization - in terms of both individual and collective users

With regard to the first category, the NMM itself is only helpful in external substantiation of internal performance. The NMM can obviously be useful for the third category, the study of communications facilities. Some experimental work has already been conducted and more is being planned. As far as the fourth category, namely user behavior, is concerned a data base is currently being analyzed with respect to the mental work expended by users in conjunction with particular kinds of commands (see next section).

Most important for this report is of course the effort on measuring interactive network-based services. In particular, one locally available system is being measured and a data base for analysis is thereby being accumulated. While it is impossible to present at this time a realistic example which would demonstrate all of the performance measurement tools described in this report, a sample set of measured and derived data can be included.

Consistent with our advocated approach that the user (or evaluator) should be able to select those measures which are of particular interest, the sample data reflect such a selection. While the data and the experiment are still incomplete, it is possible to present some statistics which are indicative of results obtainable. These statistics are in themselves valuable, because of the scarcity of data describing the workload imposed and the service rendered by an interactive computer system.

---

* Mention of a commercial product is for identification purposes only. No evaluation should be inferred.

First, a brief description of the experimental conditions. The NMM is connected to one dial-in port of the NBS Univac 1108* computer operating under Exec 8. Connection is made as shown in Figure 2b to the telephone number at the beginning of the 30 character per second rotary. There are other dial-in users as well as on-site and remote batch users. Thus these statistics represent a sampling under normal work conditions.

A sample of 69 user system sessions is involved. Summary data for that sample are displayed in Figure 4. They are listed in the following order, with each category representing several of the measures previously described in Section 2.1:

1. Time-based measures - See Table 1

2. Length-based measures - See Table 2

3. Rates - See Table 3

Figure 4 presents only the cumulative versions of each of the selected measures, across the 69 sessions. Results for the same measures can of course be produced separately for each session.

Line utilization data of particular pertinence to the determination of required communication capacity, as well as related costs (See Section 3.2), can also be generated. The cumulative version for the 69 sessions is shown in Figure 5.

Finally, it can be very useful to the analyst to have easy access to measured results through various plotted or diagrammed means of data presentation, perhaps displayed on a CRT screen. Accordingly, we can produce histograms for each of the various above-selected measures. Examples of the cumulative forms of the time-based, length based and rate measures are included in Appendix F as Parts A, B, and C respectively.

2.3 Approach to Usability

It was stated in Section 2.1.1, with regard to measures of time, that little work has been done to interpret the meanings of the user-contributed time components in an interactive session. Does user delay or "think" time (see Table 1) reflect anything about the nature of the command language used, including the lengths of the pertinent user messages (see Table 2)? Is user transmit time (Table 1) affected by message length and complexity while user delay is perhaps partly influenced by immediately preceding system

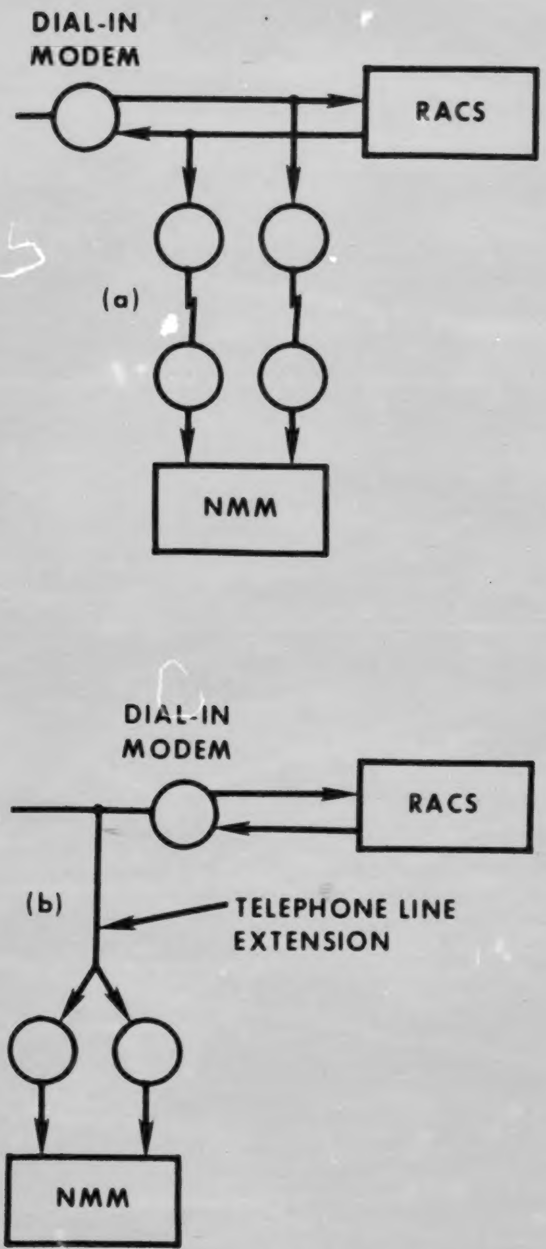


Figure 2. Remote T-Connection

(a) Connection to logic levels using four modems
(b) Connection to analog signal using two modems

```
                     SUMMARY RECORD VERSION   750523
STATISTICAL MOMENTS ARE BASED ON GROUPED DATA
ON EQUAL CLASS INTERVALS. THESE MOMENTS ARE BIASED
AND ARE NOT CORRECTED FOR GROUPING.
 RESPONSE-STIMULUS DELAY (THINK) TIME (IN SECONDS)
 TOTAL NUMBER OF OCCURRENCES=  2943 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  3026 (ABSOLUTE)
0 OCCURRENCE(S) BELOW THE MINIMUM,  83 OCCURRENCE(S) ABOVE THE MAXIMUM
MEAN=        4.7   STANDARD DEVIATION=        8.5
MEDIAN=      1.7   90%=       11.9   95%=      21.7
 STIMULUS TRANSMIT TIME (IN SECONDS)
 TOTAL NUMBER OF OCCURRENCES=  3322 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  3369 (ABSOLUTE)
0 OCCURRENCE(S) BELOW THE MINIMUM,  47 OCCURRENCE(S) ABOVE THE MAXIMUM
MEAN=        5.3   STANDARD DEVIATION=        6.8
MEDIAN=      3.0   90%=       13.7   95%=      18.8
 STIMULUS-ACKNOWLEDGEMENT DELAY TIME (IN SECONDS)
 TOTAL NUMBER OF OCCURRENCES=  3010 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  3014 (ABSOLUTE)
0 OCCURRENCE(S) BELOW THE MINIMUM,  4 OCCURRENCE(S) ABOVE THE MAXIMUM
MEAN=    .086   STANDARD DEVIATION=        0.5
MEDIAN=      0.1   90%=        0.1   95%=       0.1
 ACKNOWLEDGEMENT TRANSMIT TIME (IN SECONDS)
 TOTAL NUMBER OF OCCURRENCES=  3014
MEAN=        0.6   STANDARD DEVIATION=        1.2
MEDIAN=      0.2   90%=        1.4   95%=       2.7
 ACKNOWLEDGEMENT-RESPONSE TIME (IN SECONDS)
 TOTAL NUMBER OF OCCURRENCES=  2359 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  2402 (ABSOLUTE)
0 OCCURRENCE(S) BELOW THE MINIMUM,  43 OCCURRENCE(S) ABOVE THE MAXIMUM
MEAN=        0.4   STANDARD DEVIATION=        1.7
MEDIAN=      0.2   90%=        0.5   95%=       0.5
 RESPONSE TRANSMIT TIME (IN SECONDS)
 TOTAL NUMBER OF OCCURRENCES=  2349 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  2402 (ABSOLUTE)
0 OCCURRENCE(S) BELOW THE MINIMUM,  53 OCCURRENCE(S) ABOVE THE MAXIMUM
MEAN=        3.2   STANDARD DEVIATION=        5.4
MEDIAN=      1.5   90%=       10.3   95%=      15.1
 STIMULUS-RESPONSE DELAY TIME (IN SECONDS)
 TOTAL NUMBER OF OCCURRENCES=  2358 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  2402 (ABSOLUTE)
0 OCCURRENCE(S) BELOW THE MINIMUM,  44 OCCURRENCE(S) ABOVE THE MAXIMUM
MEAN=        1.2   STANDARD DEVIATION=        2.1
MEDIAN=      0.7   90%=        2.6   95%=       4.8
 STIMULUS INTER-ARRIVAL TIME (IN SECONDS)
 TOTAL NUMBER OF OCCURRENCES=  2994 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  3026 (ABSOLUTE)
0 OCCURRENCE(S) BELOW THE MINIMUM,  32 OCCURRENCE(S) ABOVE THE MAXIMUM
MEAN=       21.1   STANDARD DEVIATION=       37.8
MEDIAN=     10.1   90%=       40.7   95%=      65.7
```

Figure 4 Summary Data: Time-Based Measures

```
 STIMULUS CHARACTER COUNT
 TOTAL NUMBER OF OCCURRENCES=  3094
MEAN=        11.6   STANDARD DEVIATION=       12.1
MEDIAN=       8.6   90%=       23.8   95%=      33.4
 ACKNOWLEDGEMENT CHARACTER COUNT
 TOTAL NUMBER OF OCCURRENCES=  3014
MEAN=         3.2   STANDARD DEVIATION=        2.9
MEDIAN=       3.0   90%=        4.4   95%=       9.3
 RESPONSE CHARACTER COUNT
 TOTAL NUMBER OF OCCURRENCES=  2290 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  2402 (ABSOLUTE)
0 OCCURRENCE(S) BELOW THE MINIMUM,  112 OCCURRENCE(S) ABOVE THE MAXIMUM
MEAN=        56.9   STANDARD DEVIATION=       77.3
MEDIAN=      34.1   90%=      123.0   95%=     279.8
```

Figure 4 Summary Data: Length-Based Measures

```
 STIMULUS TRANSMISSION RATE (IN CHARS PER SECOND)
 TOTAL NUMBER OF OCCURRENCES=  3060 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  3094 (ABSOLUTE)
34 OCCURRENCE(S) BELOW THE MINIMUM,  0 OCCURRENCE(S) ABOVE THE MAXIMUM
MEAN=         5.5   STANDARD DEVIATION=        8.0
MEDIAN=       3.6   90%=        9.5   95%=      29.1
 ACKNOWLEDGEMENT TRANSMISSION RATE (IN CHARS PER SEC)
 TOTAL NUMBER OF OCCURRENCES=  2990 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  3014 (ABSOLUTE)
24 OCCURRENCE(S) BELOW THE MINIMUM,  0 OCCURRENCE(S) ABOVE THE MAXIMUM
MEAN=        16.4   STANDARD DEVIATION=       10.4
MEDIAN=      19.3   90%=       30.5   95%=      30.7
 RESPONSE TRANSMISSION RATE (IN CHARS PER SECOND)
 TOTAL NUMBER OF OCCURRENCES=  2387 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  2402 (ABSOLUTE)
15 OCCURRENCE(S) BELOW THE MINIMUM,  0 OCCURRENCE(S) ABOVE THE MAXIMUM
MEAN=        24.8   STANDARD DEVIATION=        5.0
MEDIAN=      26.3   90%=       30.6   95%=      30.8
```

Figure 4   Summary Data: Rates

```
LINE UTILIZATION STATISTICS     Version 750523

     300.0 Bits Per Second Transmission Line Rate

                                        Number      % of Total   % of Subtotal
Total Number of Characters              307415        100
          User                          36838          11           100
               Printing                 32623          10            88
               Non-Printing             4215            1            12
          System                        270577         89           100
               Printing                 243070         10            89
               Non-Printing             27507          10            11

Subtotal # of Printing Chars            275693         89           100
          User                          32623          10            11
          System                        243070         79            89

Subtotal # of Non-Prting Chars          31722          11           100
          User                          4215            1            13
          System                        27507          10            87

Line Utilization (Mean)
During Transmission Only (Burstiness)
          User                     7.3%
          Acknowledgement         17.9%
          Response                59.7%

Line Utilization (Mean)
          User                     3.9%
          System                  46.7%
```

Figure 5. Line Utilization for Sample Sessions

response length (Table 2)?

If we define network usability in a specialized manner to represent the human effort, or "mental work," required by a user for understanding and using alternative network/system command repertories, then the above indicated measures may be very helpful for its determination.

This section briefly presents an approach to accessing and comparing network usability based on that concept of mental work [Treu, 1975]. This is an important area which is presently not well understood; further research is indicated. First, a conceptual reference space will [illegible] developed. Then, some particular considerations in its analysis will be indicated.

2.3.1 Action Primitives

An action primitive is taken to be the core or essence of a computer-aided action conceived by a human user. After stripping the corresponding natural language command, as it might in fact be expressed by the user, of its embellishments due to grammatical structure, extraneous words, etc., and regardless of the format and content of the command after translation to a computer-recognizable form, certain basic elements and their appropriate interrelationships must be conveyed. The elements or constituents of an action primitive can simply be defined by means of the following sequence:

1. Action Verb
2. Action Qualifier(s)
3. Object(s) of the Action
4. Object Qualifier(s)

No matter what (meaningful) action is conceived in the user's mind, the nature and object(s) of the action as well as any associated qualifiers must be produced. This represents the minimal amount of mental work required of the user, preparatory to translating the desired action into language which makes sense to the system. Whether or not the user then has to explicitly indicate each of the above-listed components depends of course on the network or computer within that network involved. In any case, the conceptualized elements of each desired action must be transformed into a form which is syntactically and semantically meaningful to the system.

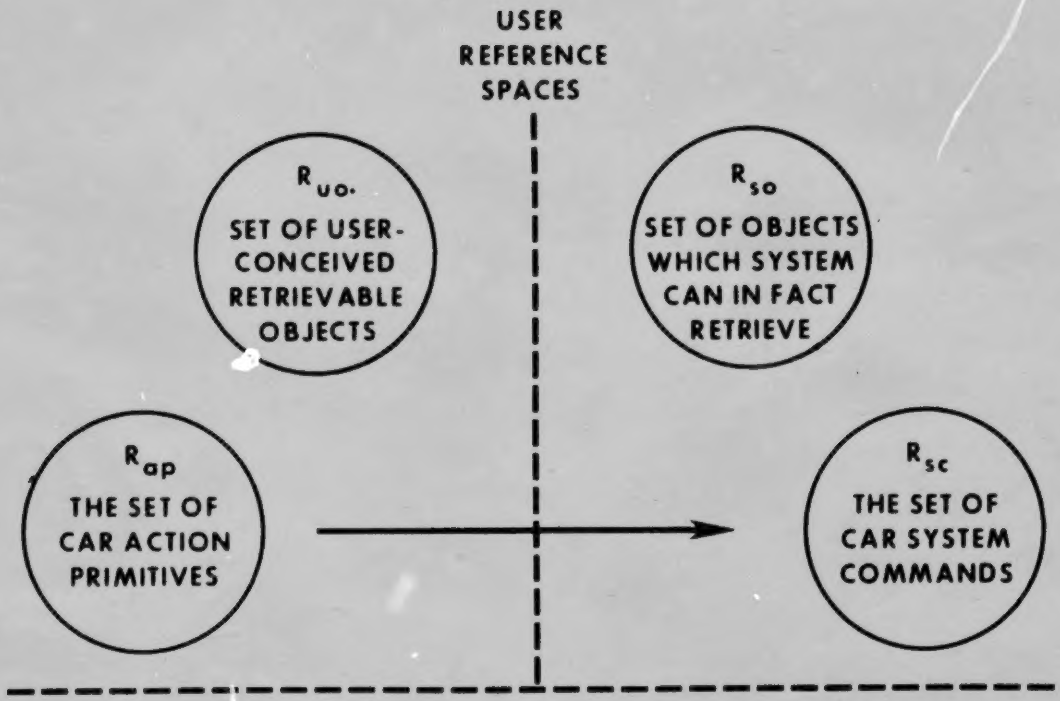
A considerable variety of computer-aided tasks is possible. They include computer-aided programming, editing, retrieval, design and many more. Some tasks also pertain to the more administrative, network-facilitated identification of and access to available resources. The conceptual reference spaces which must be constructed and utilized by users are quite different depending on which task is undertaken. For some (e.g., editing), it is very detailed and manipulative, involving many different entities and associated actions; for some (e.g., programming), it dwells more at the functional level, such as in commands for program assembly or compilation; for others (e.g., retrieval with Boolean strategy; or design using a graphics screen), considerable mental associations and conceptualizations must precede a command to search for or design something.

Nevertheless, in spite of such distinctions, it is possible to characterize all types of user-computer interaction by means of one or more of the "action primitives" defined above.

2.3.2 Transformation to Command Language

Given different types of computer-aided or network-facilitated tasks and corresponding differences in conceptual reference spaces, the action primitives relating to a particular type of task constitute a set. The result is a set of editing action primitives, a set of retrieval action primitives, and so forth. For purposes of this report the retrieval option is selected as exemplary. Its choice is based on the fact that the National Bureau of Standards is presently employing the Network Measurement Machine (see Section 2.2.1.2) for collecting data on the usability of the MEDLINE retrieval system.

For the retrieval application, the objects of on-line searching (and perhaps also storing) as conceived by the user are those to which man has become accustomed or conditioned over many years of conventional library design and use. Various representatives of articles, reports, books, etc., including their bibliographic citations, subject headings, titles, abstracts and so on, have been available in the library and should (probably) be available on-line. Hence, the comparison of sets R_{uo} and R_{so} comes into question (see Figure 6). The transformation from the set of retrieval action primitives to the set of usable retrieval system commands is hypothesized to have a direct relationship to the required mental work.


USER REFERENCE SPACES
R_{uo}.
SET OF USER-CONCEIVED RETRIEVABLE OBJECTS
R_{so}
SET OF OBJECTS WHICH SYSTEM CAN IN FACT RETRIEVE
R_{ap}
THE SET OF CAR ACTION PRIMITIVES
R_{sc}
THE SET OF CAR SYSTEM COMMANDS

ACTION	ACTION QUALIFIER(S)	R_{uo} OBJECT(S)	OBJECT QUALIFIER(S)
FIND DETERMINE STORE ⋮	I/O MEDIUM I/O FORMAT ⋮	BIBLIOGRAPHIC CITATIONS ABSTRACTS EXTRACTS FACTS SUBJECT TERMS OCCURRENCE FREQUENCIES ⋮	SUBJECT AREA(S) AUTHOR(S) OTHER QUALIFIERS E.G. COUNTRY, DATE, SECURITY LEVEL OF PERTINENCE NUMBER OF ⋮

Figure 6. Computer-Aided Retrieval

2.3.3 Associativity in Mental Work

If we claim that mental work is expended in the use of task-specific command language, it is only natural to wish to minimize the work required by a particular network-based system. But to be able to assess the (relative) performance of one or more systems in terms of how much mental work is imposed on the users, we must somehow be able to define and substantiate mental work. Neither its definition nor its substantiation is easily obtained. An attempt at the latter will, however, be suggested in the next section.

A definition or, rather, an explanation of required mental work can be approached in a two-pronged manner. First, on the hypothesis that it is dependent on the links or associations which the user must establish and then recall between/among information items, the previously defined action primitives and the corresponding system or network commands must be analyzed in detail with respect to their associative structures. Second, because a kind of translation or transformation from the conceptual command (represented by the action primitive) to the system/network command is involved, the actual bridging of that gap must also be characterized in terms of associative mental work.

It is possible to stratify the transformations into three types. If an action primitive can be directly translated into one system command, with either explicit or implicit correspondence between their respective elements, the transformation is one-to-one. If, however, a single action primitive requires the selection and issuance of two or more system commands to accomplish the intended objective, then the transformation is considered to be one-to-many. Third, the converse can exist whenever a single system command actually satisfies a composite or concatenation of action primitives.

An interesting possibility is the comparison of two different command languages designed to carry out the same network-facilitated or computer-aided task, such as retrieval. If we could agree on a set of retrieval action primitives, then analysis of each of the two languages should enable a determination of the types of transformation (i.e., one-to-one, one-to-many, many-to-one) applicable for certain subsets of the command languages respectively.

It is recognized that the above-indicated procedure is idealistic and based on a number of assumptions. Many system commands as currently implemented may, for example, be very difficult to sort out and relate to a neat set of action primitives. But that could in itself become useful information, especially if it is possible to somehow corroborate the user's level of mental effort involved.

2.3.4 Appropriate Measurements

How can "mental work" possibly be measured when psychologists and physiologists still have enormous problems in comprehending the nature and functioning of the human mind? As has happened before, we turn to external measurements, observations, and any indications which might confirm/refute hypotheses about what is going on inside.

The above-indicated user delay and transmit times are of particular interest, given our concern about required mental work. If the time from last system character to last user character is called user reaction time, then the following outlines an approximate profile of the constituents of that time:

User reaction time includes:

(1) User delay time (see Table 1) involving
system response reading and assimilation and conceptualization and command recall and selection

(2) User transmit time involving
command typing and proper/correct sequencing

The duration of each constituent is of course variable, depending on such factors as length and complexity of system message, ease of conceptualizing the next required action, ease of recalling or selecting the appropriate command, complexity and length of that command, and user's reading and typing speeds. How then could only those time segments, which are a cumulative reflection of required mental work in system command use, be separated out?

The user transmit time and the number of characters in that transmission are already being collected, as was discussed in Section 2.1. Hence a relationship between time and both pure command length and command syntactic complexity can be sought. The complexity factor will probably predominate the required mental load, unless the user is a very poor typist and thus severely affected by command length.

The most difficult problem results from the need to partition currently collected user delay times into fractions which are chargeable (primarily) to system response-related effort and next command-related work respectively. It obviously would not be fair to consider their sum as representative of the effort for conceptualizing and selecting the next system command.

Two approaches to this problem present themselves. Analogous to taking user typing speed into account in analysis of user transmit time, it may be possible to apply some proportionality factor to user delay time based on the transmission time of and number of characters/words in the previous system message. Since these data are already being obtained, and assuming that a good testable rationale for determination of a proportionality (or some other) factor is worked out, this approach appears promising.

The other approach would regrettably detract from the unobtrusiveness of dialogue monitoring by directly involving the user. By means of a simply constructed on-line indicator device used in conjunction with the interactive terminal (and the NMM), the experienced user would be asked to merely push a specific button to indicate completion of previous message reading and assimilation and thus commencement to the next task at hand. Although this procedure is subject to the usual difficulties with human inconsistencies and misunderstandings, it nevertheless is potentially useful when adequately experienced persons are available who can afford conscientious attempts at following their own thought processes.

This section on a certain aspect of computer/network usability, namely the interactive command language involved, has highlighted the applicability and interpretation of only a small subset of the measures presented in Section 2.1. But, in so doing, it shows how some currently collectible data, which tend to be largely ignored or avoided, can have a meaningful, user-oriented influence on performance evaluations and perhaps consequent procurement actions.

3. COST CONSIDERATIONS

This section complements the preceding performance considerations by examining cost factors. Certainly the cost of using a computer communications network, if not the primary criterion, is usually high on the list of important considerations upon which a selection or procurement is based. This section describes the important factors and pertinent cost models to be considered and outlines specific approaches to obtaining the necessary data to apply against the cost factors.

3.1 Costing Alternatives

The costs incurred by using a computer communications network may be based on a variety of factors and rationales. However, judging from what has been and is being actually done in charging for such services, that variety has not been great. Furthermore, it has tended to be far more oriented to costs derivable from expenditures of various "internal" computer and communication resources than to costs of real "external" performance indicators visible to and explainable by the average human user.

In view of that implied criticism, the following paragraphs outline three alternative approaches to arriving at service costs. While all three may encompass, in various ways and to varying degrees, costs of both the computer and communication resources involved, the latter are given separate, detailed coverage in Section 3.2.

3.1.1 Traditional Factors

In the past, the cost of using a computer communications service has normally been calculated on the basis of some locally established (i.e., accounted for) rates applied to all or some subset of the following types of factors respectively:

1. CPU time expended

2. Core storage area occupied

3. I/O operations required

4. Peripheral storage maintained

5. Communication channel time/capacity needed

The exact resulting cost depends on the particular algorithm, incorporating the above factors, which is employed by the computer facility in question. The algorithm may in fact only produce some locally defined number of "units" of computer utilization. These units are in turn assigned a certain monetary value. An example of this is an algorithm which is currently in use at a major university computer center:

$$[(\text{No. of Disk and Tape Operations}/120{,}000) + (\text{Connect Time Hours}/36) + 1.9/3\,(\log_2 \text{core} + 2)^{*}\ \text{CPU time} + .005]$$

The above clearly demonstrates how much the traditional methods for charging computer users have been oriented to system-internal factors to which few users may be able to relate easily. Only the connect time in the above formula is a factor which is externally observable and verifiable by the user. This is especially true when the core units and also disk/tape operations are given very special and highly technical interpretations which can in themselves be quite ambiguous.

Although it is probably unrealistic to expect that the traditional cost factors will be totally replaced in the near future by more meaningful and understandable factors, it is nevertheless useful to consider and pursue alternative approaches.

3.1.2 User-Oriented Factors

If the various measures which are derived in Section 2.1, and which are rendered usable by means of the measurement tools and techniques described in Section 2.2, are in fact representative of computer/communication service characteristics deemed important by the user, then it makes a lot of sense to consider corresponding cost factors. That is to say, the various levels or ranges of measured values attributable to a user's interaction with the network could be charged for at appropriately established rates.

Realizing of course that different levels of service require different quantities/types of internal system resources (i.e., CPU cycles, core space, etc.), it should be possible to carry out a reasonable and equitable mapping or allocation of these into the externally observable service characteristics. Hence, the previously outlined traditional factors, while still present in the background, could be replaced by the following example set of user-oriented factors:

1. Level of responsiveness

2. Level of throughput

3. Level of reliability

4. Level of verbosity

5. Length of session

An algorithm involving such factors at specified charging rates, perhaps modified by certain negative factors such as "level of user idleness," could then produce the total cost of the interactive session. Although we recognize that this approach still sounds rather idealistic and requires substantial research and development to bring it to fruition, the performance measurement methodology presented in this report is consistent with it and is likely to contribute to it.

3.1.3 Hybrid Approach

Perhaps a compromise or hybrid approach to the traditional and purely user-oriented cost factors is more immediately promising in view of our presently operational external measurement tools and techniques.

We can characterize the factors to which costs are attributable as follows:

1. Those resulting from "dynamic" (computer and communications) network performance, including
    a. CPU time expended
    b. Core storage occupied
    c. I/O operations carried out
    d. Communication channel capacity utilized
    e. Others (depending on local charging algorithm)
2. Those resulting from "static" or flat charges for (computer communications) performance or availability, including
    a. Total connect time
    b. Semi-permanent storage
    c. I/O facilities monopolized
    d. Communication line monopolized
    e. Terminal employed
    f. Others
3. Those resulting from requiring a human user, who is being serviced by the performing system, to also

perform in partnership with it,

a. Either in "active" state (i.e. reading, preparing for next command, typing, etc.)

b. Or in "inactive" state (i.e. waiting for system acknowledgement, response, completed printout of response, etc.)

4. Those resulting from other, more peripheral but contributing sources (e.g. staff assistance, documentation, etc.)

The last of these categories must be dealt with in a specialized manner or perhaps lumped into an overhead figure. Of greater interest to us are the other categories. The first two can both be regarded as traditional. The static factors, if applied, are fairly easily calculated as well as understood and generally verifiable by the user. For the dynamic factors, the usual internal system accounting procedures can be employed. As previously stated, the various aspects of communication like utilization, especially those which are monitorable by the NMM, will be addressed separately in the next section.

The third above-outlined category really brings in some user-oriented cost factors for which data are obtainable by means of the measurement methodology described in the report. If both the active and inactive user states can be ascertained through the monitoring and analysis of the various time constituents of the user-system interaction, then we need only an hourly wage or salary to be able to calculate the total user cost (active and inactive) for purposes of contrasting it against, or possibly deducting part of it from the total system cost (dynamic and static).

The above is only one example of a possible hybrid costing approach, involving both internal and certain external cost factors. It is clear that various other subsets of the measures defined in Section 2.1 could be selected and employed instead, depending on desired focus and also orientation of both system management as well as user.

3.2 Communications Costs

An earlier report [Blanc, 1973] analyzed packaged computer communications offerings (e.g., value-added networks or VAN's) in some detail with the objective of arriving at cost approximations for networking. This section, without progressing to that level of detail, will concentrate on the methodology rather than the actual costs.

Consistent with the overall theme, our user is a potential customer of a packaged communications offering, or at least desires to compare the cost of using a packaged offering with that of building his own network. The intent is to utilize cost factors from proposed tariffs, and to identify and discuss candidate techniques for arriving at the necessary data to make these factors useful to the potential user for the purposes of competitive comparisons.

3.2.1 Typical Cost Factors

Value-added network tariffs have been quite explicit in separately identifying charges relative to services rendered, with the exception of distance dependence. This is a clear and certainly a commendable attempt to adjust communications costs to accurately reflect the utilization of communications facilities and services, rather than basing costs on the leasing or purchase of raw circuit capacity. To do so, of course, requires a communications facility flexible enough to accommodate a variety of transmission speeds, efficient handling of peak traffic loads, interaction of dissimilar terminals and computers, and asymmetric traffic loads. This allows for an overall averaging effect from the service provider's viewpoint for purposes of design, raw facility procurement (circuits, modems, switches), and implementation, while still providing services, largely on a "pay-as-you-go" basis to individual users.

Typical cost factors for a value-added network are the following:

1. Packet Charge - A kilopacket is 1,000 packets. Each packet contains up to 1,024 bits, or 128 characters. For batch traffic, packets are normally full. For transaction-oriented applications, each transaction typically requires two packets: one for inquiry and one for response. A discount applies nights and weekends. Packet charges are independent of distance.

2. Port Charge for Leased Access Line - This monthly charge is applicable to both host computers and terminals and varies depending on desired Kbps line capacity or range thereof. In addition, a one-time installation charge is involved.

3. Dial-in Port Charge - Depending on desired bps rate (e.g., 1800, 2400, or 4800), a corresponding hourly

charge may be quoted. This charge is then applied in proportion to the actual length of the call.

Actual cost factors may differ significantly among vendors of VAN services. Some also include separate charges for interfacing customer-owned host computers to the communications network. For examples, charges are applied for the computer interface unit which is the special hardware channel that interconnects the customer owned host computer to a network communications processor; for the network control program which is a host resident software module that allows the host to communicate with the network; and for the connectivity modes which refer to the number of physical connects between the customer owned host computer and the network communications processor owned by the value-added network.

3.2.2 Actual Line Utilization

In order to arrive at reasonable traffic estimates, it is necessary to be familiar with existing theories, models, and measures for line utilization. It is not sufficient to know only the numbers and speeds of terminals since line utilization for a terminal is only a fraction of the maximum possible. The maximum possible line utilization is equivalent to the speed of the terminal. However, during a typical connection period, a terminal is transmitting at that speed only a portion of the time due to such factors as user think time, system response and slow typing. Because the values of these factors can be determined, as described in Section 2, it is possible to arrive at accurate line utilization estimates. These estimates can then be integrated over all terminals attached to a network, to arrive at total traffic calculations.

Unfortunately, very little work has been done in this area. The data stream model [Jackson and Stubbs, 1969] does attack this problem in modeling the communications process between an interactive user and a multiaccess computer system. The model defines a character burst as a string of characters related to each other through the property of interarrival times below a certain threshold. The communication process is then separated into four functional parts:

1. user send time (the total amount of time during which user characters are being transmitted);

2. computer send time (the total amount of time during which computer characters are being transmitted);

3. user delay (the sum of all inactive periods during user burst segments);

4. computer delay (the sum of all inactive periods during computer burst segments).

Data are gathered on a large number of calls to each of several multiaccess computer systems and applied against the model for the purpose of determining the attributes of holding time. Relationships are derived between that holding time and computer delays, user characteristics, and computer send time. Above-listed measures clearly correspond to certain measures defined in Section 2, although they are labelled differently.

The analysis performed aptly identifies some of the important characteristics of the user/system interaction. The quantization of the various delays, both user introduced and system introduced, can lead to the conclusion that it would be desirable in using communication services to pay for data transmitted rather than holding time. The need for asymmetric channels is demonstrated in the derived quantity relationship between average number of characters sent by the computer and those sent by the user, the former being an order of magnitude greater than the latter. The bursty nature of traffic in the user/system interaction is also demonstrated and measured. These burst characteristics will be important to the discussion of blocking or packeting data for transmission through a network.

The relationships derived between holding times and computer delays, user characteristics, and computer send time are supported by data which show that similar relationships exist between amount and nature of traffic-generated and computer delays, user characteristics, and computer send time. This indicates, that, for a user interested in deriving traffic estimates reflective of his own terminal activities, it would be insufficient to use the data of the data stream model or of any other generalized study. It is instead necessary to collect similar or expanded statistics for a specific user, user community, or control group represented by that community. The intricacies of characterizing the user/system interaction and of collecting and analyzing data applicable to specific system and user populations is realized. The paper reporting the data stream model [Jackson and Stubbs, 1969] identifies the following implications:

1. new data gathering procedures and equipment are needed;

2. data analysis procedures must be capable of handling very large quantities of data; and

3. legal, ethical, and business requirements related to communications and computing privacy must be satisfied.

The first two implications will be further discussed. With regard to the last, it is assumed that the owner or manager of the measured facility is also the originator of the request and the recipient of the measurement results.

3.2.3 Data Collection Techniques

One technique for collecting data on the user system interaction is that represented by the Network Measurement Machine (NMM) described in Section 2. This approach is representative of a class of approaches whereby data are collected externally to the network and at the user interface. As was previously discussed, the NMM can connect between a user at (an) asynchronous terminal(s) and the terminal interface to a network (e.g., concentration). This connection is depicted in Figure 7.

The data on individual time-tagged characters and their sources are collected on tape and that tape serves as input to a set of analysis routines which provide the actual statistics for the user system interactions. Figures 7 and 8 illustrate such a set of statistics. From the standpoint of using these statistics to estimate traffic generating potential of a specific user, it becomes a fairly straightforward calculation to derive average line utilization over a given time period of both system and user. The specific timing information, as a second order benefit, leads to intercharacter arrival (transmit and receive) statistics which can be used to quantitatively describe the burst characteristics of the user/system interaction.

From the standpoint of the user or potential user of a packaged computer communications offering, those traffic data represent the missing (and possibly most significant) set of statistics useful in determining the cost of using the VAN. This determination allows for the objective comparison of competitive offerings and can lead to a comparison of the packaged offerings and the "build your own network approach." The significance of the traffic data and its associated cost factors are supposed from examining the cost factors of proposed VAN's. A cost analysis [Blanc, 1974] makes this importance even more obvious, even though the traffic statistics used in the report are hypothetical,

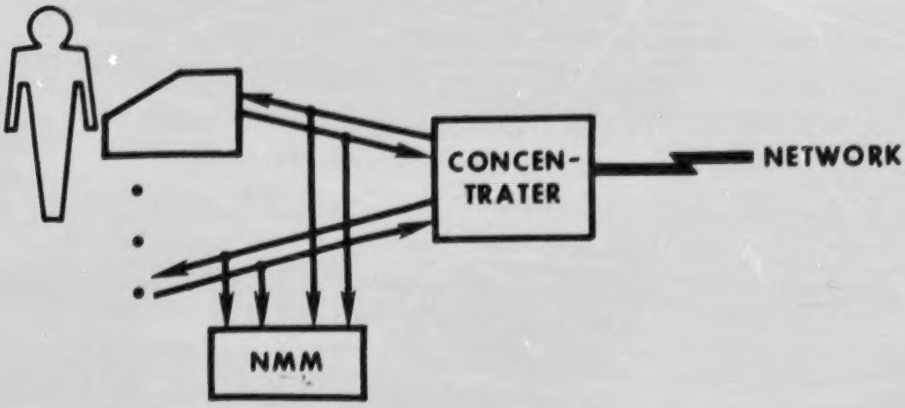


Figure 7.

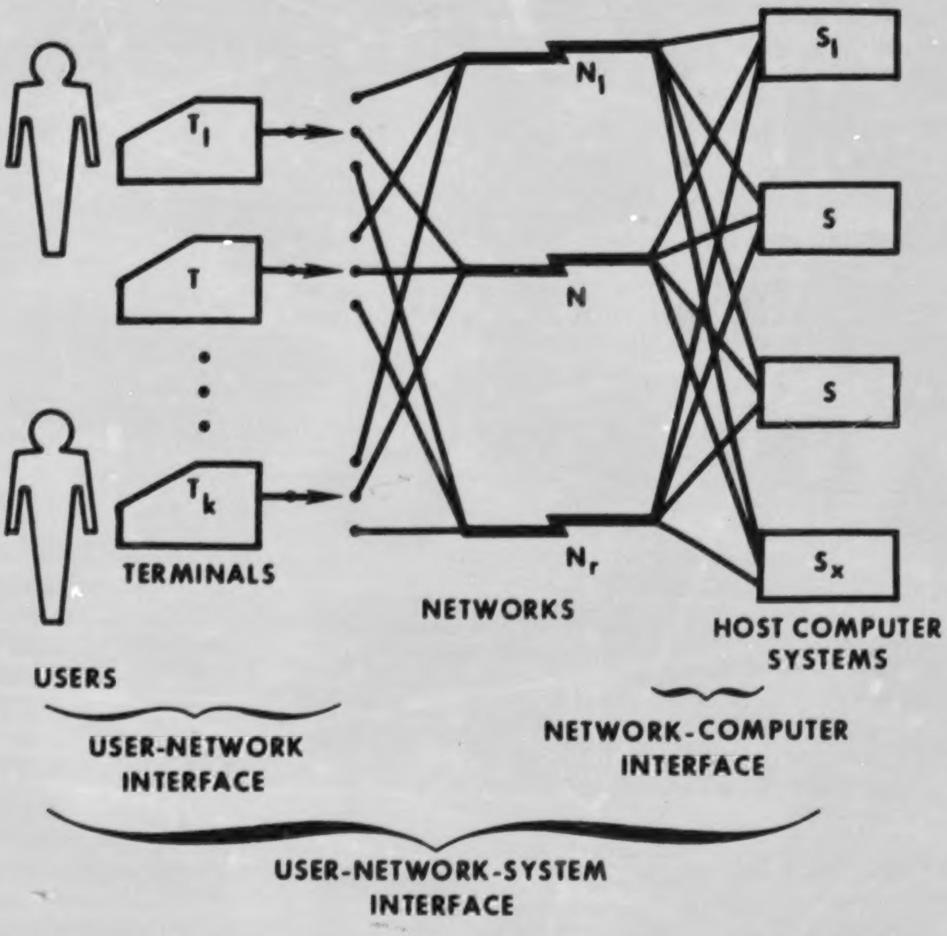


Figure 8. Alternative User Interfaces

based on the somewhat dated work of Jackson and Stubbs [1969] and on only preliminary data of the NMM.

Similarly the burst characteristics of the user/system interaction may prove to be important. At this time it is not altogether clear how the proposed VAN's will block or "packetize" data. But it appears that rates will be based on the number of blocks or packets transmitted rather than the number of characters. It is conceivable that the number of characters per packet (up to a pre-defined maximum) will be determined by the burst size of the user input and system output character strings. Using the simplest case, the burst size may be assumed equal to the line length. Then in the proposed offerings, a user transmitting short lines will pay more than a user transmitting long lines, even though the total number of characters in the user system interaction are identical. The quantification of "the more" can be derived from the statistics collected and generated about the user/system interaction by the NMM.

3.2.4 Data Traffic Characterization

It should not be concluded that these kinds of statistics are useful only to the prospective user of a packaged communications offering. On the contrary, the communication's system designer and implementer should be equally, if not even more, concerned. During this last year the National Academy of Engineering Panel on Telecommunications Research underscored this need by stating that there is a "lack of a good characterization of the statistics of data traffic". Communication system implementers have gone about installing large systems, in spite of the lack of a data base of good statistics describing user generated data traffic. The source of the problem has been the lack of suitable techniques and measurement devices to capture such data.

Statistics describing data traffic would be useful in the overall network layout considerations, as well as in the determination of the capacities of the individual network components. These components range from circuits, modems, multiplexers, and concentrators, to the sophisticated devices that serve as host interfaces, terminal interfaces, message switching computers, and even the host computers and their associated front-end devices. Not knowing traffic statistics can result in network bottlenecks due to insufficient capacity in specific network components, as well as, wasteful deployment of high capacity components, through allowing for unrealistic worst-case situations. Traffic statistics describing the user/system or user/network interaction, such as those generated by the

Network Measurement Machine or similar devices as they develop, should prove particularly useful in this application area.

In addition to the data traffic as described by line utilization, the burst characteristics of the user/system interaction are also of interest. Since large and medium sized computer communication networks block data from interactive terminals at the terminal interface for transmission on the network, knowing data traffic burst characteristics would be useful in determining the optimal blocking factors. Once again a good data base of statistics describing the burst characteristics of the user/system interaction would contribute a great deal to more effective and efficient communication system design. As previously discussed, devices like the Network Measurement Machine, which capture the appropriate data at the user interface, can be utilized in developing such a data base.

4. SUMMARY AND CONCLUSIONS

In the introductory section, three main objectives for this report were outlined. Instead of simply repeating them here, it may be helpful to characterize them differently and thereby not only summarize but perhaps also further clarify our approach.

Our predominant concern has been with identifying the means and methods for assisting the present or prospective users and/or procurers of network-based computer services in assessments of network performance and cost. This user/procurer orientation has led us to develop and advocate "external" measurement techniques. The diagrammatic portrayal in Figure 8 may serve to interpret that term properly, within the context of this report.

The figure indicates how one or more users may have interest in and (potential) access to one or more computer system alternatives, by means of one or more (packaged) communications networks offerings. External measurement is then taken to encompass both the immediate user-network interface, with respect to what the user experiences directly at the terminal, and also the more distant interface between user and the serving computer system itself. The latter type of interface clearly subsumes measurement of the performance and cost of the intermediary communications network.

Given that portrayal of our measurement activities, we were faced with having to identify or develop suitable measures which upon their application, could serve to satisfy the criteria on whether one system or network is performing well or better than another. Similarly, questions about (comparative) network costs were to be answerable. Performance- related measures were presented and discussed extensively in Section 2.1; cost considerations and their relationships to certain performance measures were dealt with in Section 3.

But measures are clearly not enough. To render them useful, they require that appropriate measurement tools and techniques be designed and implemented. Such tools and techniques were described in Section 2.2, both with respect to measurement applications involving real users as well as situations in which highly controlled test procedures with simulated users are demanded or preferred.

Our measurement orientation, together with the described measures and tools and techniques, must furthermore be integrated into one or more effective measurement methodologies. Methodology is, in a sense, intertwined throughout the sections of this report. It is reflected in the tabulations of measures in Section 2.1, the measurement tool of Section 2.2, the specialized approach to usability in Section 2.3, and the comparison of network cost factors of Section 3. Nevertheless, although the constituents are already present, a neatly packaged methodological product, perhaps in the form of a well-structured user's guide to network performance measurement, is not yet available.

There have been several identifiable results produced as a consequence of the work covered in this report. They include:

(1) Identification of the relevant basic measures; primarily response time, breadth of service provided and processing capability;

(2) Further demonstration of the feasibility and usefulness of employing a hardware/software tool, such as the Network Measurement System, for purposes of "external" measurement of computer communications network performance thereby eliminating the need for separate consideration of communications facility and host computer;

(3) Recognition that the complexity of the network performance issue required a versatile capability

which can measure a variety of events and produce meaningful resorts derived from these measurements; consistent with the types of objectively collectible data (by means of the NMS), the establishment of a comprehensive framework of measures which can, in fact, be meaningfully interpreted and applied both in conjunction with performance and cost determinations;

(4) The explanation, modeling, and application of the measurement tools and techniques with the objectives of serving real, pragmatic needs, such as evaluation and procurement of computer communications networks or network devices; and

(5) The initiation and development, throughout the above, of a very desirable user-oriented methodology of bringing the tools and techniques for computer network performance measurement out to where the users are and can understand them, as opposed to continuing to leave those people at the mercy of highly specialized and "internalized" performance measurement approaches.

Finally, because the work described in this report has provided the investigators with considerable insights into the means and methods whereby a more user-oriented approach to performance and cost measurement is possible, a specific recommendation is in order.

Both pre- and post-procurement purposes are potentially well-served by the measurement methodology presented. However, while significant progress has been made, much work remains to be done. In particular it is clear, and therefore recommended, that considerably more testing and subsequent refinement of the described methodology be carried out. This should precede and then lead to the development and publication of a comprehensive and understandable methodological guide to network performance measurement.

BIBLIOGRAPHY AND REFERENCES

Adams, J. and Cohen, L. [1969]. "Time-sharing vs. Instant Batch Processing." Computers and Automation, pp. 33-34.

Arbuckle, R. A. [1966]. "Computer Analysis and Throughput Evaluation," Computers and Automation, Vol. 15, No. 1, January 1966, pp. 12 - 15.

ASCII - 1968. USA Standard Code for Information Interchange, American National Standards Institute, X3.4-1968.

Abrams, M. D. and Cotton, I. W. [1975]. The Service Concept Applied to Computer Networks, National Bureau of Standards Technical Note 880, August 1975.

Abrams, M. D., Lindamood, G. E. and Pyke, T. N., Jr. [1973]. "Measuring and Modeling Man-Computer Interaction," Proc. of the 1st SIGME Symposium on Measurement and Evaluation (ACM), February 1973, pp. 136-142.

Abrams, M. D. [1974]. "A New Approach to Performance Evaluation of Computer Networks," Proc. of the 1974 Symposium COMPUTER NETWORKS: Trends and Applications (IEEE), May 1974, pp. 15-20.

Boies, S. J. [1974]. "User Behavior on an Interactive Computer System," IBM Systems Journal, No. 1, 1974, pp. 2-18.

Boehm, B. W., Seven, M. J. and Watson, R. A. [1971]. "Interactive Problem-Solving-An Experimental Study of "Lockout Effects," Proc. Spring Joint Computer Conference, 1971, pp. 205-210.

Buchholz, W. [1969]. "A Synthetic Job for Measuring System Performance," IBM System Journal, Vol. 8, No. 4, 1969, pp. 309 - 318.

Calingaert, P. [1967]. "System Performance Evaluation: Survey and Appraisal." Communications of the ACM, 10:1 pp.12-17.

Carbonell, J. R., Elkind, J. I. and Nickerson, R. S. [1968]. "On the Psychological Importance of Time in a Time Sharing System," Human Factors, Vol. 10, No. 2, 1968, pp. 135-142.

Carlisle, J. H. [1970]. "Comparing Behavior at Various Computer Display Consoles in Time-shared Legal Information." RAND Corporation Report P-4448.

Cease, D. R. [1970]. TSS/360 Experience Using One Approach to Addressing Interactive System Performance, TR 21.400, IBM Systems Development Division, November 1970.

Computer and Business Equipment Manufacturers Association (CBEMA) [1975]. Use of Remote Terminal Emulation in ADP Procurement to Validate the Performance of Large Terminal Oriented Systems, (cover letter) March 24, 1975.

Crothers, C. G. [1974]. Workload Determination and Representation for On-Line Computer Systems, ESD-TR-74-54, The MITRE Corporation, Bedford, MA (NTIS AD-779 818), January 1974.

De Meis, W. M. and Weizer, N. [1969]. "Measurement and Analysis of a Demand Paging Time Sharing System," Proc. 24th ACM National Conference, 1969, pp. 201 - 216.

Dexter, A. S. [1973]. "Analysis of a Commercial Time-Shared Environment," Infor, Vol. II, No. 1, February 1973, pp. 41-49.

Doherty, W. J., Thompson, C. H. and Boies, S. J. [1972]. "An Analysis of Interactive System Usage with Respect to Software, Linguistic, and Scheduling Attributes," IBM Research, RC3914, June 30, 1972, p. 25.

Electronic Industries Association [1969]. EIA Standard RS-232-C, Interface Between Data Terminal Equipment and Data Communications Equipment Employing Serial Binary Data Interchange, August 1969.

Erikson, W. J. [1966]. "A Pilot Study of Interactive Versus Noninteractive Debugging." System Development Corporation, TN-3296.

Federal Property Management Regulations, Amendment E-131 [1973]. "Restrictions on the Use of Simulation in the ADPE Procurement Process", Subpart 101-32.408-1, The Federal Register, 38 F. R. 22227, August 17, 1973.

Ferrari, D. [1972]. "Workload Characterization and Selection in Computer Performance Measurement," Computer, July/August 1972, pp. 18 - 24.

FIPS PUB 31 [1974]. *Guidelines for Automatic Data Processing Physical Security and Risk Management*, Federal Information Processing Standards Publication 31, June 1974.

FIPS PUB 42 [1975]. *Preliminary Guidelines for Benchmarking ADP Systems in the Competitive Procurement Environment.* Federal Information Processing Standards Publication 42, September 1975.

Gold, M. M. [1967]. "Methodology for Evaluating Time-shared Computer Usage." Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Alfred P. Sloan School of Management.

Gold, M. M. [1969]. "Timesharing and Batch-processing: an Experimental Comparison of their Values in a Problem-solving Situation." *Communications of the ACM*, 12:5 pp. 249-259.

Grant, E. E. and Sackman, H. [1967]. "An Exploratory Investigation of Programmer Performance Under On-line and Off-line Conditions." *IEEE Transactions on Human Factors in Electronics*, pp. 33-48.

Greenbaum, H. J. [1969]. *A Simulator of Multiple Interactive Users to Drive a Time-Shared Computer System*, MAC-TR-58, (Thesis) Massachusetts Institute of Technology, Project MAC, January 1969.

Grenander, U. [1973]. *An Introduction to Compumetrics*, Brown University Report (NTIS AS 766 470), July 1973.

Grubb, D. S. and Cotton, I. W. [1975]. *Criteria for the Performance Evaluation of Data Communications Services for Computer Networks.* NBS Technical Note 882, September 1975.

Hillegass, J. R. [1969]. "Systematic Techniques for Computer Evaluation and Selection." *Management Services*, pp. 35-38.

Holdsworth, D., Robinson, G. W., and Wells, M. [1973]. "A Multi-Terminal Benchmark," *Software - Practice and Experience*, Vol. 3, No. 1, January - March 1973, pp. 43-59.

Jackson, P. E. and Stubbs, C. D. [1969]. "A Study of Multiaccess Computer Communications," *Proc. Spring Joint Computer Conference, 1969*, pp. 491 - 504.

Joseph, G. W. [1972]. The FORTRAN Frequency Analyzer as a Data Gathering Aid for Computer Simulation, ESD-TR-72-159, The Rand Corporation, March 1972.

Joslin, E. O. [1965]. "Application Benchmarks: The Key to Meaningful Computer Evaluation," Proc. 20th ACM National Conference, 1965, pp. 27 - 37.

Joslin, E. O. [1971]. "Costing the System Design Alternatives." Data Management , pp. 23-27.

Joslin, E. O. and Aiken, J. J. [1966]. "The Validity of Basing Computer Selections on Benchmark Results," Computers and Automation, Vol. 15, No. 1, January 1966, pp. 22 - 23.

Joslin, E. O. and Mullin, M.J. [1964]. "Cost-value Technique for Evaluation of Computer System Proposals," Proc. Spring Joint Computer Conference, pp. 367-381.

Joslin, E. O. [1968]. Computer Selection, Addison-Wesley Publishing Company, Inc.

Joslin, E. O., ed. [1971]. Analysis, Design and Selection of Computer Systems, College Readings, Inc., Arlington, Va.

Kamerman, A. [1969]. Establishing and Measuring Conversational Performance Objectives for Time Sharing Systems, TR 53.015, IBM Systems Development Division (Yorktown Heights, N.Y.), June 1969.

Karp, P.M. [1973]. "Origin, Development and Current Status of the ARPA Network," Proc. Compcon 73, pp. 25-30.

Knuth, D. E. [1971]. "An Empirical Study of FORTRAN Programs," Software - Practice and Experience, Vol. 1, No.2, April - June 1971, pp. 105 - 134.

Lampson, B. W. "A Critique of 'An Exploratory Investigation of Programmer Performance Under On-line and Off-line Conditions.'" IEEE Transactions on Human Factors in Electronic, pp. 48-51.

Lassettre, E.R. and Scherr, A.L. [1972]. "Modelling and Performance of the OS/360 Time-Sharing Option (TSO)," In: Statistical Computer Performance Evaluation (Ed. by W. Freiberger), Academic Press, pp. 57-72.

Lyon, G. and Stillman, R. [1974]. A FORTRAN Analyzer, NBS Technical Note 849, October 1974.

Mamrak, S. A. [1975]. Comparative Response Times of Time-Sharing Systems on the ARPA Network, Ph.D. Dissertation, University of Illinois, UIUCDCS-R-75-722, May 1975.

Miller, E. F., Jr. [1972]. "Bibliography on Techniques of Computer Performance Analysis," Computer, September/October 1972, pp. 35-47.

Miller, R. B. [1968]. "Response Time in Man-Computer Conversational Transactions," Proc. Fall Joint Computer Conference, 1968, Vol. 33, Part 1, pp. 267-277.

Mitre Corp. [1973-1975]. "Remote Terminal Emulator (Design Verification Model), Bedford, Mass., Technical Report MTR-2677, Volumes as cited below:

Vol. 1, " - Introduction and Summary," February 1975, 76p. (By D. L. James and D. W. Lambert).

Vol. 2, " - Data Structures, Scenario Instructions and Commands," November 1973, 79p. (By D. L. James).

Vol. 3, " - Macro Preprocessor," November 1973, 47p. (By M. L. Dolley).

Vol. 4, " - Scenario Assembler," August 1973, 83p. (By Tobyanne Paster).

Vol. 5, " - Scenario Interpreter," April 1974, 227p. (By D. L. James)

Vol. 7, " - Data Reduction Program," April 1974, 129p. (By M. C. Sarajian)

Vol. 9, " - Support Software and Hardware Diagnostics," December 1973, 48p. (By M. C. Sarajian).

Morgan, D. E., Banks, W., Colvin, W., and Sutton, D. [1974]. A Performance Measurement System for Computer Networks," Proc. IFIP Congress 1974.

Nelson, E. A. [1967]. "Management Handbook for the Estimation of Computer Programming Costs." System Development Corporation, Technical Memorandum No. 3225.

Neuman, A. J., Lucas, B. G., Walker, J. C., and Fife, D. W. [1974]. A Technical Guide to Computer-Communications Interface Standards, NBS Technical Note 843, August 1974.

Rosenthal, R., Rippy, D. E. and Wood, H. M. [1975]. The Network Measurement Machine-A Data Collection Device for Measuring the Performance and Utilization of Computer Networks, National Bureau of Standards Technical Note 912, April 1976.

Rosenthal, R. and Watkins, S. W. [1974]. "Automated Access to Network Resources; A Network Access Machine," Proc. of the 1974 Symposium COMPUTER NETWORKS: Trends and Applications, May 1974, pp. 47-50.

Sackman, H. [1970]. Man-Computer Problem Solving. Princeton, N. J.: Auerbach Publishers Inc.

Sackman, H. [1972]. Mass Information Utilities and Social Excellence. Princeton, N. J.: Auerbach Publishers Inc.

Saltzer, J. H. and Gintell, J. W. [1970]. "The Instrumentation of Multics," Communications of the ACM, Vol. 13, No. 8, August 1970, pp. 495-500.

Schatzoff, M., Tsao, R., and Wiig, R. [1967]. "An Experimental Comparison of Time Sharing and Batch Processing." CACM, pp. 261-65.

Schwemm, R. E. [1972]. "Experience Gained in the Development and Use of TSS," Proc. Spring Joint Computer Conference, 1972, pp. 559 - 569.

Sharpe, W. F. [1969]. The Economics of Computers. Columbia University Press: New York, (RAND Report R-463-PR).

Silverman, H. F. and Yue, P. C. [1973]. "Response Time Characterization of an Information Retrieval System," IBM J. Research and Development, Vol. 17, No. 5, September 1973, pp. 394-403.

Smith, J. M. [1968]. "A Review and Comparison of Certain Methods of Computer Performance Evaluation," Computer Bulletin, 12:1, pp. 13-18.

Smith, L. B. [1967]. "A Comparison of Batch Processing and Instant Turnaround," CACM, pp. 495-500.

Streeter, D. N. [1972]. "Cost-Benefit Evaluation of Scientific Computing Services," IBM Systems Journal No. 3, 1972, pp. 219-233.

Subcommittee on Priorities and Economy in the Government of the Joint Economic Committee [1972]. "American Productivity: Key to Economic Strength and National Survival," July 1972. Systems in the Competitive Procurement Environment, Federal Information Processing Standards Publication 42, September 1975.

Tesdata Systems Corporation [1973]. Load Generator System General Information Manual, (undated) received January 1973.

Treu, S. [1975]. "Interactive Command Language Design Based on Required Mental Work," International Journal of Man-Machine Studies, Vol. 7, No. 1.

Walkowicz, J. [1974]. Benchmarking and Workload Definition: A Selected Bibliography with Abstracts, NBS Special Publication 405, November 1974.

Watkins, S. W. and Abrams, M. D. [1976]. Data Treatment and Analysis for the Network Measurement System, National Bureau of Standards Technical Note 897, March 1976.

Watson, R. [1971]. Computer Performance Analysis: Applications of Accounting Data, R-573-NASA/PR, The Rand Corporation, May 1971.

Vareha, A. [1972]. SLIN, unpublished notes and private conversation, 1972.

Wood, D. C. and Forman, E. H. [1971]. "Throughput Measurement Using a Synthetic Job Stream," Proc. Fall Joint Computer Conference, , pp. 51 - 56.

Appendix A

Delimiting Measurements

In order to measure and analyze events it is necessary to define the coordinates which will be employed to locate and identify these events. In this report, the fundamental event of interest is the occurrence of a character in context. The most commonly used coordinates are time and character ordinal numbers. Time might be local time-of-day or elapsed time since the beginning of the measurement unit. Because time is a continuous variable, there are associated questions of precision and accuracy. Characters are counted in discrete integers which are in some ways more easily handled.

Assuming that a selection of coordinates has been made so that it is possible to identify an event, we must decide which events are of sufficient interest. First, let us define the range in which we are interested. Although the single character is a measurable quantity with which various instantaneous values may be associated, we shall consider groups or sets of characters as our lower limit. Actually, since the entire user-system network data flow to be measured is describable as a set of characters, we must be more specific.

A sequence of characters typed by the human computer user is called a "message" or "human input" or "stimulus." A sequence of characters coming from the computer is called "computer output" or "response." In most interactive situations stimuli and responses are associated, although not necessarily on a one-to-one basis. The set of one stimulus and one or more associated response(s) is called a "transaction" or "message group." For highly interactive full-duplex communication, as for example with TENEX, it may be necessary to introduce sub-stimuli and sub-responses in a message group. A set of message groups closely connected in time constitutes a "conversation" or "session" between man and computer. Sets of sessions may be defined in terms of single descriptors such as name of person or computer system, or they may be defined as logical combinations of such descriptors.

As far as temporal data are concerned, the computer output has been divided into two subsets. The first subset called "acknowledgement," contains no information relevant to the content of the stimulus, but rather is relevant only to the existence of the stimulus. In the most simple case the acknowledgement consists of a line feed (LF) to position the paper and a few other control characters to provide timing delay, unlock the keyboard, etc. A heading of the form "the computer says" is an example of a more complex acknowledgement. That part of the computer output which is not acknowledgement is the "(true) response."

Given these definitions of sets of characters, it is easy to define points such as the beginning of the stimulus, the end (or the last character) of the stimulus, the beginning of the response, etc. Good reasons can be offered for measuring between most pairs of such points. The selection of pairs of measurement points is dependent on the objective of the measurement activity. The only problem is that different investigators tend to use the same name for different point pairs and the associated measurements. This is further amplified in Subsection 2.1.3 with regard to "response time."

Appendix B

Versions of Response Time

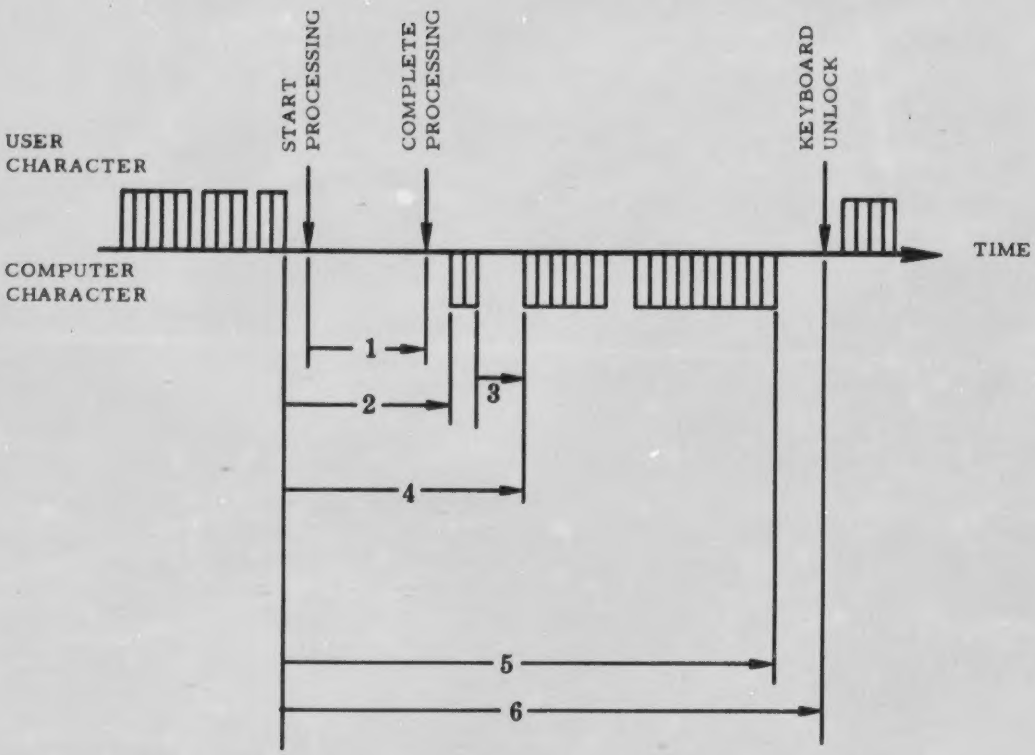
In his extensive discussion of response time, Miller [1968], presents seventeen different stimuli to which there must be responses. While many of these stimulus-response situations reflect psychological considerations beyond the scope of this report, his points deserve very serious consideration. Several of his measures are incorporated in more service oriented measures, to which we now turn our attention.

Figure 9 displays six different selections of elapsed time which are in use as measures of computer system response. As an introduction to the precise and formal definitions which follow, let us briefly characterize these different definitions.

Silverman and Yue [1973] employ the name "Inquiry Response Time" to refer to the amount of processing time required for response to an inquiry stimulus. Since the system they measured operated with queuing on both input and output, processing could not begin until some time after the arrival of the last character of the stimulus. Similarly, the first character of the computer output response could not appear until some time after processing was completed. While these internal measures may be easily measured by software and are (non-deterministically) related to observable events, their utility appears to be limited to parametric studies within a single server computer system. Comparison among network-based servers requires measures based on externally meaningful and measurable events.

A very common measure is the elapsed time from the last character of user stimulus input until the first character of computer output, indicated as (2) in the figure. A closely related measure is based on the recognition that some number of initial characters from the computer may not convey any meaningful response information to the user (see Subsection 2.1.1); therefore, elapsed time is measured from the last character of the user stimulus until the first meaningful character of the computer output. This measure is indicated as (4) in the figure.

The acknowledgement serves a psychological function of reassuring the user of the computer's continuing ability to serve him; it may also serve a practical function of advancing the print position. With the introduction of the acknowledgement as a substate of computer output, measure (2) takes on the meaning of the elapsed time from the end of



1. Inquiry Response Time [Silverman, 1973]
2. Response Time, System Responsiveness [Kamerman, 1969]
   Stimulus-Acknowledgement Delay [Abrams & Cotton, 1975]
3. Acknowledgement-Response Delay [Abrams & Cotton, 1975]
4. Stimulus-Response Delay [Abrams & Cotton, 1975]
5. Turnaround Time, Response Time [Boehm, 1971]
6. System Response Time [Boies, 1974]

VARIOUS MEASURES OF SYSTEM RESPONSE

Figure 9

the stimulus to the beginning of the acknowledgement. Another delay, indicated as (3), is introduced to measure the elapsed time from the end of the acknowledgement to the beginning of the response.

Still another measure which is intuitively appealing is the elapsed time from the last character of the user stimulus to the last character of the computer output. Indicated as (5) in the figure, this measure incorporates the verbosity and transmission efficiency of the server into a single measure which is especially valuable for inter-server comparison. A closely related measure is the elapsed time from the end of the stimulus until the computer is able to accept another stimulus. This latter state, made obvious by keyboard unlock or printing of a "prompt character," is identified as (6).

Appendix C

Modes of Use of the Network Measurement Machine

For terminals local to the NMM and without appropriate modems, the NMM may be hardwired to the terminal. This connection is very useful, in fact, in that dial-out capability of the NMM may be employed to provide the otherwise absent modems. As shown in Figure 10, all connections to the NMM are brought through a patch panel, thereby making it possible to connect terminals to various resources, including the NMM, with ease. The more common method of connecting terminals through the NMM is the dial-in dial-out capability. The NMM is equipped with two sets of modems, each of which has its own rotary of telephone numbers. Instead of dialing the number which he/she normally calls for computer service, the user dials the NMM. The terminal speed and code are identified by typing the letter "E," to which the NMM responds with a salutation and a request for the telephone number which normally would have been dialed in order to obtain computer service. An automatic calling unit connected to the second set of modems dials the call. Success or failure to establish data communication is announced to the user.

The method of connecting the user to the computer while measuring the data traffic deserves further elaboration. The first stage, which we call initial connection, simply has the user connected to the NMM as shown schematically in Figure 11a. When the connection is made to the remote computer, the situation changes to that shown in Figure 11b. We rejected a store-and-forward connection through the NMM because of its potential for perturbing the system to be measured. It would cause a fixed delay of two character-times for transmission in each direction plus possible additional delay due to processing within the NMM. Instead we have chosen the more fail-safe connection wherein the "input" and "output" modems are cross-connected with control and receive-data transmit-data lines interchanged as necessary. The lines providing input to the NMM are the receive-data, the transmit-data, and the carrier detect lines of one of the modems. The typing by the user of the telephone number to which he wishes to be connected can be generalized. The names of computer systems can be substituted for telephone numbers if a table is kept in the NMM. If the computer system is local, the dial-out operation may be replaced by a direct connection. This connection is carried out when the user wishes to be connected to our TIP onto the ARPANET [Karp, 1973]. Many of the functions of computer service acquisition through a network currently under study at NBS in the Network Access

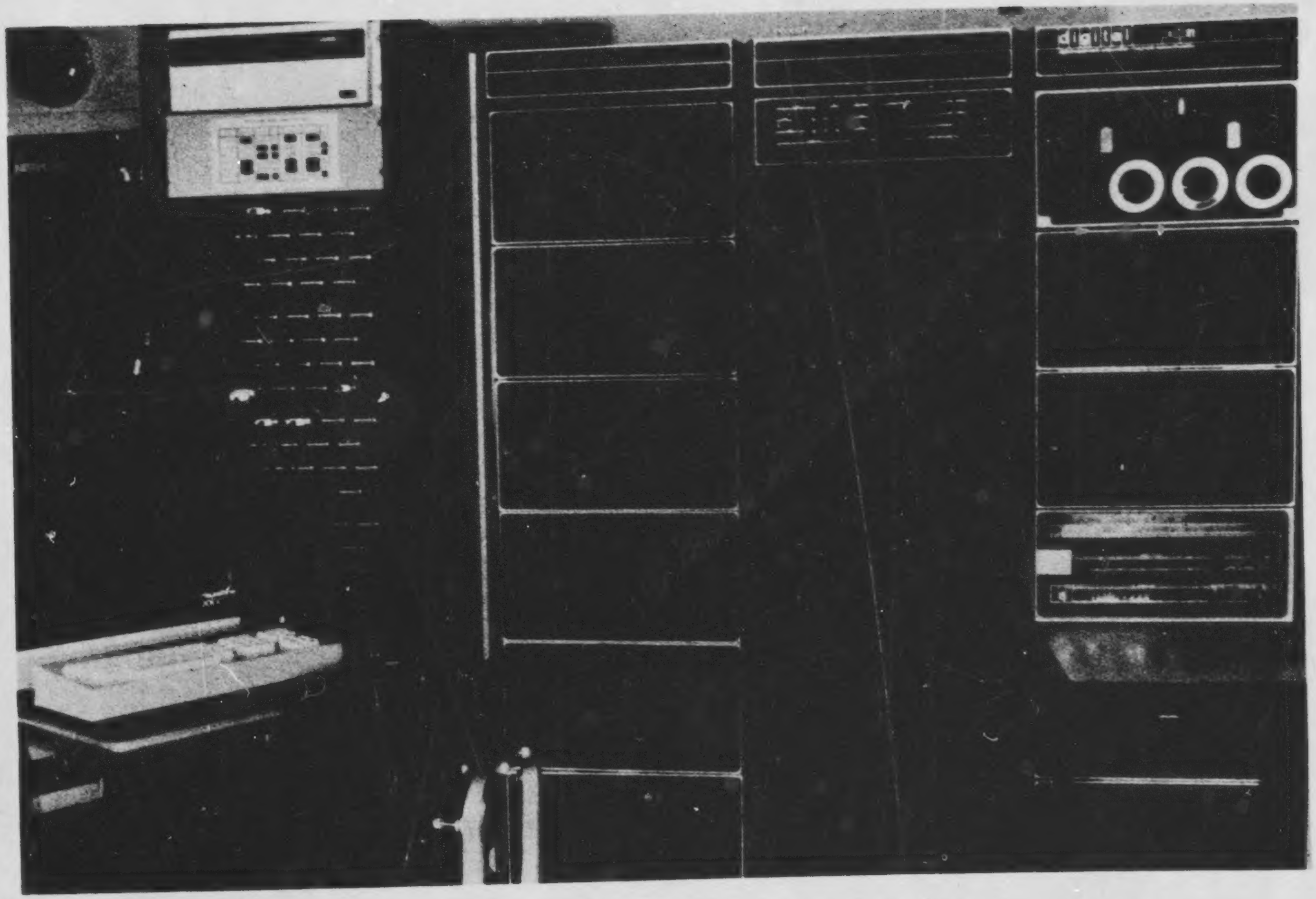


Figure 10. The Network Measurement Machine

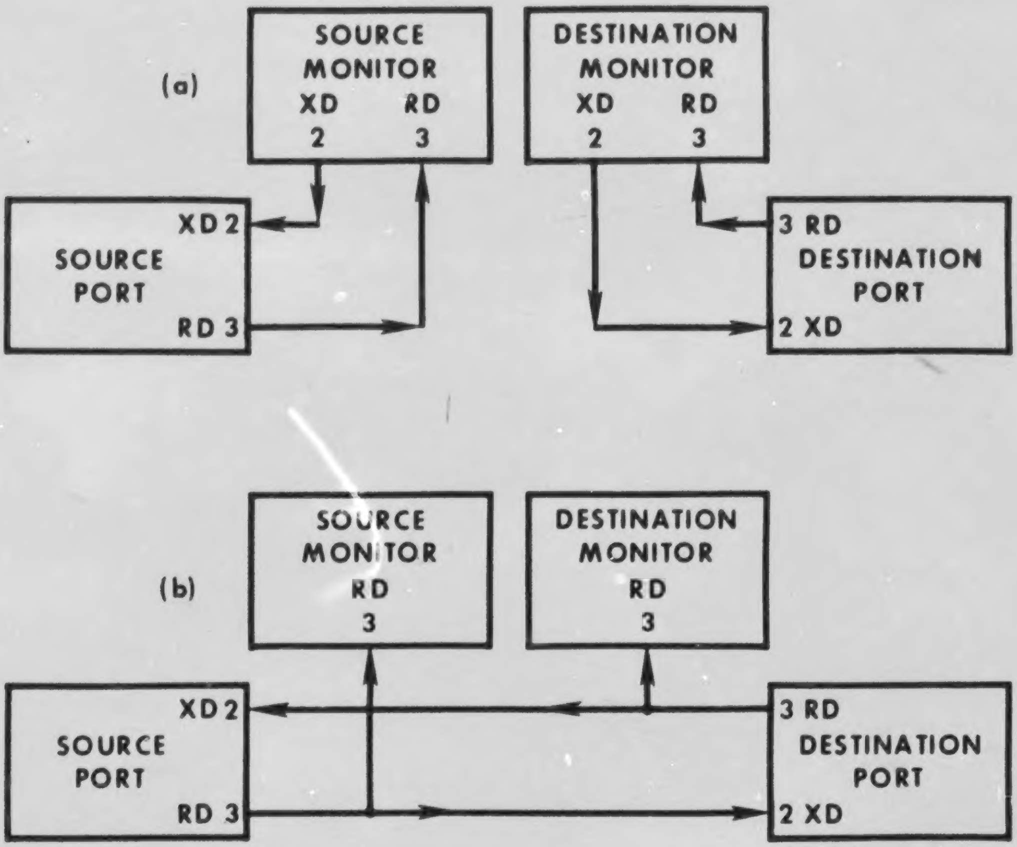


Figure 11. Connection States of the Network Measurement Machine

(a) NMM connected to user (source) and destination
(b) Source and destination directly connected

Machine (NMM) [Rosenthal and Watkins, 1974] could be combined with the NMM in a very powerful and useful system.

The NMM may also be used to gather information of particular interest to the producer of computer services; in this mode the NMM is connected to the communications interface of the host computer. (For certain types of measurement discussed below, there may also be connection to the user's terminal.) Three methods are available for this connection. If the NMM is local to the host computer, connection may be made to the RS232 logic levels in the same manner as was done for the user terminal. This method is extendible to the situation in which the NMM is remote from the host by using modems at each end of a common carrier line as shown in Figure 2. So far, we have used two modems at each end of two private lines in order to accomplish our objective. This configuration employs one pair of modems and a private line to make the transmit data signal available to the NMM and the other pair for the receive data signal. We have specified but not yet implemented a "dual-simplex" modem which could be produced by modifying commercial modems that would permit us to use one dial-up line to transfer both signals. Our third method, shown in Figure 2b, is to have an "extension" of the analog dial-in line at the host computer brought to the NMM and connected to a pair of modems modified similarly to that described above through a special telephone company-supplied high impedance amplifier.

As long as the NMM is connected to the host's interface, it is in a position to record the data communications. Unlike the connection to the terminal, a logic signal is not necessarily generated when one user terminates his conversation with the host and when another user begins. This lack of hardware determination of end-of-conversation must be compensated for in the analysis software as described below.

The two preceding fundamental modes, connection to the user's terminal and connection to the system's interface, may be combined for special measurements. By way of illustration, two experimental designs will be briefly sketched. In order to measure the comparative effect of a network on the service rendered to a user, we have devised and employed the configuration shown in Figure 3 wherein two different networks are used to connect to the same serving host. After login has been accomplished at the separate terminals, the switch is thrown so that the transmit data line on one "master" terminal controls the transmit data line on all the modems. In this way, identical conditions, except for differences in the transmission facilities

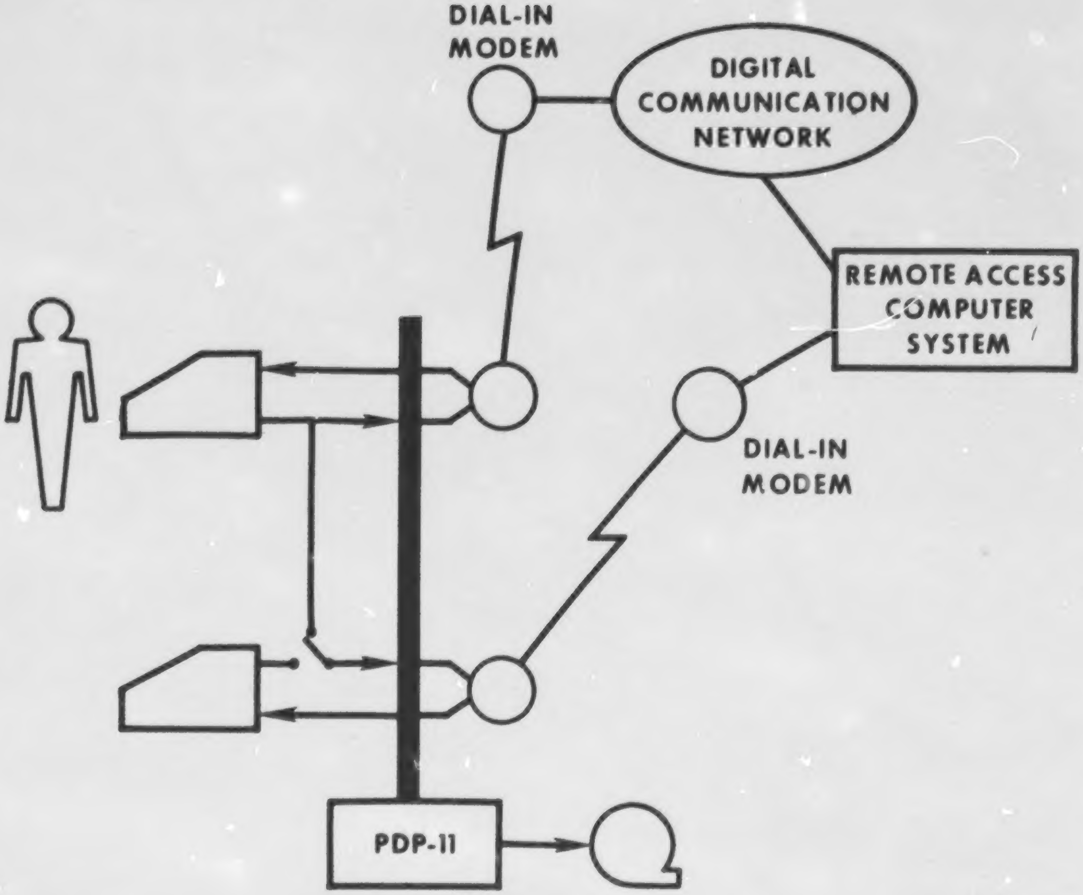


Figure 3. Network Measurement Machine

themselves, are assured. In another experiment being planned, the NMM is connected to both user terminal and host interface for the one single conversation. This simultaneous recording assures timing synchronization so that the absolute effect of the transmission network may be determined.

Appendix D

Data Treatment in the Network Measurement Systems

Practical implementations of the conceptual models discussed in section 2.2.1.1 occur in the data analysis programs. The data measured by the NMM is not analyzed in real time. That capability might be valuable for adjustment of system operation parameters, such as the number of conversational jobs permitted. There is no inherent reason why this could not be done; we were simply not interested in such a function. Our operational mode is to dump the data measured by the NMM onto tape for transportation to a large time sharing computer system for subsequent analysis.

The intended use of these models has been to derive information about: (a) user demands for service from the remote access system, (b) communications facility utilization, and (c) remote computer service (e.g., meaningful response time statistics). The last of these is, of course, of primary interest in this report.

Using the SAR model, a variety of measurements concerning the conversation can be made. These measurements fall into two broad classifications: those concerned with character count, and those involving elapsed time. The character count is the number of characters occurring in each message segment. (See Table 2).

Because transmission and delay times are measured by the elapsed time between two events, numerous measurements may be made by simply altering the definition of an event. Because of the multiplicity of definitions to be introduced in the next three subsections, the names of events explicitly identify the interval being measured. This policy of using full phrase names rather than abbreviations can alleviate some of the existing terminological confusion. (See also Figure 9).

D.1. Standard Sampling Interval

To eliminate unusual and unrepresentative occurrences, a standard sampling interval is defined to exclude what are called "outlayers." It is imperative that any pathological cases be eliminated from the data so as not to distort the statistics. An on-line user of a computer system can become distracted by and involved in an activity totally unrelated to his/her computer usage during that usage. It is also possible for a system to crash at any point during a conversation. The occurrence of these types of events produce the distorting data. To recognize the presence of these data, analyst-set upper and lower limits must be used. These limits formulate the standard sampling intervals, and data must fall within the interval to be considered in the statistics.

The acknowledgement and response delay time limits should be set high enough to assure that the system has crashed and is not just heavily loaded. The acknowledgement and response transmission time limits are based on the average character speed. If the character count divided by the transmission time is greater than a set multiple of the character speed capability of the line, the data should be discarded.

If the analyst is concerned with precise measures, the type of user/system interaction must be considered before setting the stimulus delay limit. Conversations involving dynamic user decision-making should be allowed a higher limit than those which do not. However, for the analyst interested only in conversations in general, the limit can be set independently of conversation type. The stimulus (or user) transmission limit may be based on the same type of calculation as the system response transmission, since the user's ability as a typist need not be known. In this case, it is desirable to eliminate the data reflecting a user's leaving the terminal in the midst of character transmittal to the computer system. Therefore, the limit may be set assuming worst case ability and still be adequate.

By dividing the standard sampling interval into a number of sub-intervals it is possible to characterize the distribution of the derived parameters by counting the number of occurrences of a parameter value in each of the sub-intervals, and the more accurate the representation of the distribution. Practical programming considerations limit the number of sub-intervals. It is, therefore, desirable to make the standard sampling interval as small as possible. By this criterion the upper and lower limits on the standard sampling interval should be carefully chosen to exclude uninteresting as well as anomalous values. Since there is a minor conflict between these two preceding criteria for establishing the standard sampling interval, this selection should be done very carefully.

D.2. Full Duplex -- Half Duplex

Relative to our interests, full-duplex transmission has two major effects. First, it increases network utilization by virtue of echoes. Second, it makes the demarcation of user and computer transmission segments more difficult.

In many, but not all, cases, the Network Measurement Machine record of a full-duplex conversation may be reduced to an equivalent half-duplex case. This transformation makes it possible to employ the same data reduction techniques to both transmission types. The implementation of this transformation records sufficient data concerning the increased network utilization.

Those computers which exhibit a sophisticated utilization of the full-duplex mode (e.g., TENEX), cannot have their conversation records completely algorithmically reduced to the half-duplex model, at least by the present implementation. The analyst must then decide whether to complete the transformation to the half-duplex model manually, or to represent this conversation by the full-duplex model.

Since the half-duplex network utilization mode is simplest to understand and analyze, the half-duplex model containing the salient points of this mode serves most of the purposes of our current measurements. Whenever possible, data obtained from a full-duplex conversation are transformed according to the half-duplex model.

D.3. Conversation Record

The execution of the analysis routines does not necessitate a hard copy reproduction of the conversation, but if a record is requested, the user specifies the terminal being used. The program maintains pertinent statistics (i.e., the number of rows and columns per page, and the number of rows and columns per inch) for a variety of terminals.

Because many characters have no printed representation, in implementation these characters take the form of the standard abbreviation enclosed in corner brackets. For example, a carriage return would appear as <CR>. For the sake of readability the space character is treated as a special case. If a space is the first or last character on a line, it appears as <SP>; otherwise, a blank character is printed.

Appendix E

Measurement Drivers

E.1. Measurement Driver Characteristics

Basically there are two approaches to implementing a measurement driver; it may be either internal or external to the system under test (SUT). If internal, the measurement driver program runs in a memory partition disjoint from the rest of the system and causes interrupts so as to deceive the communications device handler into believing that it is communicating with external terminals. If external, the measurement driver equipment or hardware connects to the SUT through its communications device interfaces, either locally or through a communications network, and interacts with the SUT as if it were a set of interactive terminal users. Alternatively, it may be connected to the SUT through a single high speed port. Although the internal measurement driver perturbs the system being measured and fails to completely test the extremities of the terminal handler, in practice these deficiencies are either proven to be negligible or are ignored. It is therefore reasonable to consider them as variants on the same basic theme.

The data obtained from the NMM, in addition to providing us with considerable experience, have sufficiently convinced us that these systems which are to be measured are non-deterministic. If the SUT is deterministic, the measurement driver need not examine computer output because it is known in advance. If, however, the SUT is non-deterministic, computer output must be examined to determine if the unexpected has perchance occurred. Unsolicited messages such as "WAIT," "PLEASE DELETE ALL UNNEEDED FILES," etc., are frequent. Changes in the operating system messages also exist. It is a relatively safe first-order approximation to treat a SUT as being deterministic. Occasionally the driver which assumes determinism will get out of synchronization with the SUT, producing a useless series of trivial interactions. Among the considerations as to whether the SUT is to be considered deterministic or not are the questions of repeatability and flexibility. To a certain extent, trade-offs must be made between these two objectives. Repeatability requires that, each time the simulator presents an activity (run, job) to a computer, the differences observed are due to the computer system and not to the simulator. Flexibility requires that the simulator not abort an otherwise useful activity due to an extraneous perturoation in the environment.

Each of the measurement drivers to be described employs a generalization of the benchmark called a script. The term script is purposefully borrowed from the theater. A script is the specification of the conversational dialogue between the user and the computer. A deterministic script will probably not explicitly specify the computer part of the dialogue.

When the computer is recognized to be non-deterministic, the script must contain both user and system halves of the conversation. In either case the script contains not only the characters which constitute the conversation, but also timing information describing the temporal relationship among the characters. It is important to note that a script is more than the generalization of a benchmark program. Since the script completely specifies the dialogue between man and computer, it contains many elements in addition to a source language program and a set of data. For example, a script might specify logging onto a system, creating a program to be stored in the file system, followed by compiling, linking and loading, and executing. In execution, input and output might be conducted between the program and the terminal as well as between the program and various files. A script may contain a mixture of commands at the executive command level, subcommands and other interactions with various subsystems, programs, and data all interwoven in a conversational format.

E.2. Examples of Measurement Drivers

Several measurement drivers exist, having been produced as for-sale commercial products, as proprietary system development testers and as general purpose tools. The following example selection is described for purposes of representativeness, not completeness. Furthermore, no attempt at comparison is made.

SLIN [Vareha, 1972] is a modification of the TSS/360 Supervisor providing a general purpose testing capability which allows evaluation under heavy usage without requiring large amounts of communications hardware. Any number of terminal users may be simulated up to the system limit. Scripts describing the users are input through the card reader, any number of scripts may be used during a run, and one or more simulated users may execute each script. Each card specifies the text for one line typed at the terminal plus the total minimum transaction time for that cycle of the conversational benchmark (see description of TEST/360).

Tesdata Systems Corporation is marketing the Load Generator System for the Honeywell 600-6000 systems operating under GCOS [Tesdata, 1973]; this Load Generator is an extension of work done for the Defense Intelligence Agency. The Load Generator is a batch program under GCOS which operates as a modification to the terminal communications subsystem, permitting real and simulated terminals to coexist. Up to 180 terminal users may be simulated.

TEST/360 is an internal tool used by IBM in the development of TSS and presumably TSO. Minimum operating configuration is a 360/40 which is connected to the SUT through low speed communications ports. Although references are scarce, Kamerman [1969] describes the objectives in the implementation and application of TEST/360. Cease [1970] and Schwemm [1972] discuss the application environment of TEST/360 and its user's reactions. Kamerman states that "conversational performance objectives involve the execution of a defined system load consisting of a specified number of terminals running a specified set of conversational benchmarks upon a defined system configuration. With the system so loaded each benchmark should be completed in a satisfactory length of time.... The criteria for successful achievement of the performance objectives should state the limits of acceptable elongation of the benchmark sessions.... At any point in the development of a timesharing system, one should be able to say how many terminals are supported and what they are doing."

Eight criteria are proposed for TEST/360: "a. It must interface with the system in the same manner as actual terminals and, to the system, appear indistinguishable from them; b. It must be capable of bringing the system up to a stable load level and keeping it there during the measurement period in order to eliminate starting and stopping transients from a measurement; c. It must be capable of recognizing the expected system response for each transmission and contain sufficient logic to take appropriate action from any response received; d. It must be capable of recovery from at least the ordinary kind of unexpected response, such as might result from transmission errors; e. It should be able to accept as a parameter, a "human keying rate" for each "terminal," to adjust for the difference between machine driven transmission rates and real keying rates; f. It must record and time stamp all transmissions in both directions on all lines, together with the control parameters and other such data, to allow data reductions and analysis at a later time; g. It should be capable of terminating a run, based upon cut-off criteria for that run, to avoid wasteful use of machine time; h. It should be data directed in its operation, so that not only transmission to be transmitted but also the control of the individual delays between each transaction may be specified in the conversational benchmarks, in accordance with the measurement rules." According to our information, objectives c and d either were not implemented in TEST/360 or are sufficiently difficult to use that most scripts do not incorporate these features.

The MITRE Remote Terminal Emulator (RTE) project is conducted under the sponsorship of the Air Force Directorate for Automatic Data Processing Equipment Selection and is motivated by competitive procurement of large scale EDP systems for the United States Air Force [Mitre, 1973-1975]. The emulator is separate from the SUT to which it is connected by means of data communication lines. A language has been developed in which to write scripts which are then translated into an internal form by a script assembler. The assembler scripts are then interpretively executed. Responses are logged, but not analyzed. There are data reduction programs which list in time sequence all of the records logged by the Emulator and will in addition produce lists in time sequence by device. Some summarizing and computation is also done by the data reduction program, but little emphasis is placed on this activity.

Greenbaum [1969] developed a PDP-8 based measurement driver which has been used to simulate users of CTSS and MULTICS. The simulator recognizes and verifies responses transmitted by the SUT. Scripts, encoded in a special script creation language, contain the information necessary to perform checking and verification plus think-time emulation.

In addition to the text lines that a human user would normally type at his console, the script also includes a "verifier line" used by the simulator to check and verify responses transmitted to it by the SUT. This script also includes a "verifier time limit" and a "think-time." The former allows the simulator to check for normal operation of the SUT during simulation. The latter allows the simulator to mimic the time spent in "head scratching" by a human user. The simulator provides for the capability of recognizing abnormal situations which may occur during operation. Some of these are: a. transmission errors on the data communication lines; b. system crashes of the SUT; c. abnormal SUT operation leading to unexpected responses. The simulator can detect these problems, but detection is only half of the problem. What could, or should, be done to correct these errors? Greenbaum concludes that if the system crashes, obviously the experiment should be repeated when the SUT is back in operation. Otherwise, he questions the desirability of restarting the current simulation. He reasons that the simulation is a tool to be used to provide an experimenter with a means of performing a controlled repeatable experiment on a SUT whereby he can make performance measurements on the SUT. Implementation of the restart feature puts the repeatability of such a simulation in question. He goes on to state that the single largest problem facing the experimenter is the design of his scripts.

Appendix F

Statistics

STATISTICAL MOMENTS ARE BASED ON GROUPED DATA
ON EQUAL CLASS INTERVALS. THESE MOMENTS ARE BIASED
AND ARE NOT CORRECTED FOR GROUPING.

```
           RESPONSE-STIMULUS DELAY (THINK) TIME (IN SECONDS)
 0.0 (1205) !-------------------------------------------------
 1.2 (642)  !--------------------------
 2.4 (273)  !-----------
 3.6 (156)  !------
 4.8 (97)   !---
 6.0 (92)   !---
 7.2 (75)   !---
 8.4 (49)   !-
 9.6 (30)   !-
10.8 (39)   !-
12.0 (32)   !-
13.2 (24)   !-
14.4 (16)   !-
15.6 (21)   !-
16.8 (15)   !-
18.0 (16)   !-
19.2 (9)    !-
20.4 (5)    !-
21.6 (6)    !-
22.8 (9)    !-
24.0 (6)    !-
25.2 (13)   !-
26.4 (8)    !-
27.6 (9)    !-
28.8 (6)    !-
30.0 (4)    !-
31.2 (7)    !-
32.4 (3)    !-
33.6 (3)    !-
34.8 (5)    !-
36.0 (1)    !-
37.2 (10)   !-
38.4 (2)    !-
39.6 (6)    !-
40.8 (4)    !-
42.0 (5)    !-
43.2 (2)    !-
44.4 (3)    !-
45.6 (2)    !-
46.8 (6)    !-
48.0 (2)    !-
49.2 (6)    !-
50.4 (1)    !-
51.6 (5)    !-
52.8 (3)    !-
54.0 (0)    !
55.2 (3)    !-
56.4 (4)    !-
57.6 (1)    !-
58.8 (2)    !-
FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
GROUP FREQUENCY RANGE    .  1.2
 TOTAL NUMBER OF OCCURRENCES=  2943 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  3026 (ABSOLUTE)
0 OCCURRENCE(S) BELOW THE MINIMUM,  83 OCCURRENCE(S) ABOVE THE MA
XIMUM
MEAN=        4.7   STANDARD DEVIATION=       8.5
MEDIAN=      1.7   90%=       11.9   95%=     21.7
             Part A  Histograms for Time-Based Measures
```

```
                  STIMULUS TRANSMIT TIME (IN SECONDS)
 0.0 (878)  !--------------------------------------------------
 1.0 (487)  !----------------------------
 2.0 (331)  !------------------
 3.0 (264)  !--------------
 4.0 (215)  !-----------
 5.0 (176)  !---------
 6.0 (126)  !-------
 7.0 (121)  !------
 8.0 (104)  !-----
 9.0 (87)   !----
10.0 (62)   !---
11.0 (69)   !---
12.0 (46)   !--
13.0 (39)   !--
14.0 (36)   !--
15.0 (34)   !-
16.0 (34)   !-
17.0 (29)   !-
18.0 (27)   !-
19.0 (21)   !-
20.0 (12)   !-
21.0 (12)   !-
22.0 (12)   !-
23.0 (10)   !-
24.0 (10)   !-
25.0 (4)    !-
26.0 (7)    !-
27.0 (5)    !-
28.0 (6)    !-
29.0 (5)    !-
30.0 (6)    !-
31.0 (1)    !-
32.0 (7)    !-
33.0 (4)    !-
34.0 (6)    !-
35.0 (5)    !-
36.0 (2)    !-
37.0 (5)    !-
38.0 (1)    !-
39.0 (2)    !-
40.0 (1)    !-
41.0 (1)    !-
42.0 (0)    !
43.0 (4)    !-
44.0 (0)    !
45.0 (2)    !-
46.0 (3)    !-
47.0 (2)    !-
48.0 (0)    !
49.0 (1)    !-
FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
GROUP FREQUENCY RANGE      1.0
 TOTAL NUMBER OF OCCURRENCES=  3322 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  3369 (ABSOLUTE)
0 OCCURRENCE(S) BELOW THE MINIMUM,  47 OCCURRENCE(S) ABOVE THE MA
XIMUM
MEAN=        5.3   STANDARD DEVIATION=       6.8
MEDIAN=      3.0   90%=      13.7   95%=      18.8
```

Part A (Continued)

```
                  STIMULUS-ACKNOWLEDGEMENT DELAY TIME (IN SECONDS)
        0.0 (2978) !------------------------------------------------
        0.4 (10)   !-
        0.8 (2)    !-
        1.2 (2)    !-
        1.6 (3)    !-
        2.0 (4)    !-
        2.4 (1)    !-
        2.8 (0)    !
        3.2 (1)    !-
        3.6 (1)    !-
        4.0 (1)    !-
        4.4 (1)    !-
        4.8 (0)    !
        5.2 (0)    !
        5.6 (1)    !-
        6.0 (0)    !
        6.4 (0)    !
        6.8 (0)    !
        7.2 (0)    !
        7.6 (0)    !
        8.0 (2)    !-
        8.4 (0)    !
        8.8 (0)    !
        9.2 (0)    !
        9.6 (0)    !
       10.0 (1)    !-
       10.4 (0)    !
       10.8 (0)    !
       11.2 (0)    !
       11.6 (0)    !
       12.0 (0)    !
       12.4 (0)    !
       12.8 (0)    !
       13.2 (1)    !-
       13.6 (0)    !
       14.0 (0)    !
       14.4 (0)    !
       14.8 (0)    !
       15.2 (0)    !
       15.6 (0)    !
       16.0 (1)    !-
       FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
       GROUP FREQUENCY RANGE      0.4
        TOTAL NUMBER OF OCCURRENCES=  3010 (IN RANGE)
        TOTAL NUMBER OF OCCURRENCES=  3014 (ABSOLUTE)
       0 OCCURRENCE(S) BELOW THE MINIMUM,  4 OCCURRENCE(S) ABOVE THE MAX
IMUM
       MEAN=   .086   STANDARD DEVIATION=       0.5
       MEDIAN=     0.1   90%=      0.1   95%=      0.1
                                Part A  (Continued)
```

```
              ACKNOWLEDGEMENT TRANSMIT TIME (IN SECONDS)
  0.0 (2144) !--------------------------------------------------
  0.4 (395)  !---------
  0.8 (141)  !---
  1.2 (75)   !-
  1.6 (33)   !-
  2.0 (44)   !-
  2.4 (44)   !-
  2.8 (29)   !-
  3.2 (11)   !-
  3.6 (13)   !-
  4.0 (11)   !-
  4.4 (9)    !-
  4.8 (7)    !-
  5.2 (6)    !-
  5.6 (6)    !-
  6.0 (4)    !-
  6.4 (10)   !-
  6.8 (3)    !-
  7.2 (5)    !-
  7.6 (4)    !-
  8.0 (3)    !-
  8.4 (1)    !-
  8.8 (7)    !-
  9.2 (2)    !-
  9.6 (6)    !-
 10.0 (0)    !
 10.4 (1)    !-
FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
GROUP FREQUENCY RANGE      0.4
 TOTAL NUMBER OF OCCURRENCES=  3014
MEAN=        0.6   STANDARD DEVIATION=      1.2
MEDIAN=      0.2   90%=      1.4   95%=      2.7
```

Part A (Continued)

```
                     ACKNOWLEDGEMENT-RESPONSE TIME (IN SECONDS)
        0.0 (1763) !-----------------------------------------------
        0.4 (555)  !--------------
        0.8 (0)    !
        1.2 (0)    !
        1.6 (0)    !
        2.0 (0)    !
        2.4 (0)    !
        2.8 (0)    !
        3.2 (0)    !
        3.6 (0)    !
        4.0 (0)    !
        4.4 (0)    !
        4.8 (1)    !-
        5.2 (0)    !
        5.6 (0)    !
        6.0 (0)    !
        6.4 (0)    !
        6.8 (0)    !
        7.2 (0)    !
        7.6 (0)    !
        8.0 (0)    !
        8.4 (0)    !
        8.8 (0)    !
        9.2 (0)    !
        9.6 (0)    !
       10.0 (4)    !-
       10.4 (2)    !-
       10.8 (2)    !-
       11.2 (2)    !-
       11.6 (3)    !-
       12.0 (3)    !-
       12.4 (4)    !-
       12.8 (4)    !-
       13.2 (1)    !-
       13.6 (0)    !
       14.0 (2)    !-
       14.4 (3)    !-
       14.8 (2)    !-
       15.2 (2)    !-
       15.6 (2)    !-
       16.0 (0)    !
       16.4 (2)    !-
       16.8 (0)    !
       17.2 (2)    !-
       FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
       GROUP FREQUENCY RANGE      0.4
        TOTAL NUMBER OF OCCURRENCES=  2359 (IN RANGE)
        TOTAL NUMBER OF OCCURRENCES=  2402 (ABSOLUTE)
       0 OCCURRENCE(S) BELOW THE MINIMUM,  43 OCCURRENCE(S) ABOVE THE MA
XIMUM
       MEAN=        0.4   STANDARD DEVIATION=       1.7
       MEDIAN=      0.2   90%=       0.5   95%=       0.5
```

Part A (Continued)

```
                         RESPONSE TRANSMIT TIME (IN SECONDS)
        0.0 (606)  !----------------------------------------
        0.8 (732)  !-------------------------------------------------
        1.6 (353)  !-----------------------
        2.4 (230)  !---------------
        3.2 (66)   !----
        4.0 (32)   !--
        4.8 (24)   !-
        5.6 (15)   !-
        6.4 (11)   !-
        7.2 (7)    !-
        8.0 (14)   !-
        8.8 (9)    !-
        9.6 (35)   !--
       10.4 (30)   !--
       11.2 (21)   !-
       12.0 (20)   !-
       12.8 (8)    !-
       13.6 (6)    !-
       14.4 (15)   !-
       15.2 (16)   !-
       16.0 (10)   !-
       16.8 (4)    !-
       17.6 (14)   !-
       18.4 (7)    !-
       19.2 (7)    !-
       20.0 (6)    !-
       20.8 (3)    !-
       21.6 (5)    !-
       22.4 (6)    !-
       23.2 (0)    !
       24.0 (3)    !-
       24.8 (2)    !-
       25.6 (0)    !
       26.4 (2)    !-
       27.2 (1)    !-
       28.0 (1)    !-
       28.8 (2)    !-
       29.6 (7)    !-
       30.4 (1)    !-
       31.2 (4)    !-
       32.0 (1)    !-
       32.8 (4)    !-
       33.6 (0)    !
       34.4 (2)    !-
       35.2 (0)    !
       36.0 (2)    !-
       36.8 (1)    !-
       37.6 (2)    !-
       38.4 (1)    !-
       39.2 (1)    !-
       FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
       GROUP FREQUENCY RANGE      0.8
        TOTAL NUMBER OF OCCURRENCES=  2349 (IN RANGE)
        TOTAL NUMBER OF OCCURRENCES=  2402 (ABSOLUTE)
       0 OCCURRENCE(S) BELOW THE MINIMUM,  53 OCCURRENCE(S) ABOVE THE MA
XIMUM
       MEAN=        3.2   STANDARD DEVIATION=       5.4
       MEDIAN=      1.5   90%=      10.3   95%=     15.1
```

Part A (Continued)

```
              STIMULUS-RESPONSE DELAY TIME (IN SECONDS)
 0.0 (804)  !-----------------------------------------------
 0.4 (818)  !------------------------------------------------
 0.8 (268)  !----------------
 1.2 (111)  !------
 1.6 (62)   !---
 2.0 (43)   !--
 2.4 (42)   !--
 2.8 (38)   !--
 3.2 (19)   !-
 3.6 (12)   !-
 4.0 (15)   !-
 4.4 (10)   !-
 4.8 (12)   !-
 5.2 (3)    !-
 5.6 (6)    !-
 6.0 (6)    !-
 6.4 (6)    !-
 6.8 (9)    !-
 7.2 (4)    !-
 7.6 (3)    !-
 8.0 (5)    !-
 8.4 (4)    !-
 8.8 (4)    !-
 9.2 (3)    !-
 9.6 (5)    !-
10.0 (5)    !-
10.4 (5)    !-
10.8 (2)    !-
11.2 (2)    !-
11.6 (3)    !-
12.0 (3)    !-
12.4 (3)    !-
12.8 (4)    !-
13.2 (2)    !-
13.6 (0)    !
14.0 (3)    !-
14.4 (4)    !-
14.8 (1)    !-
15.2 (2)    !-
15.6 (2)    !-
16.0 (1)    !-
16.4 (2)    !-
16.8 (0)    !
17.2 (2)    !-
FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
GROUP FREQUENCY RANGE      0.4
 TOTAL NUMBER OF OCCURRENCES=  2358 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  2402 (ABSOLUTE)
0 OCCURRENCE(S) BELOW THE MINIMUM,  44 OCCURRENCE(S) ABOVE THE MA
XIMUM
MEAN=        1.2   STANDARD DEVIATION=       2.1
MEDIAN=      0.7   90%=        2.6   95%=      4.8
                        Part A  (Continued)
```

```
                 STIMULUS INTER-ARRIVAL TIME (IN SECONDS)
   0.0 (822)  !-------------------------------------------------
   5.0 (737)  !-------------------------------------------
  10.0 (436)  !-------------------------
  15.0 (279)  !----------------
  20.0 (170)  !----------
  25.0 (100)  !-----
  30.0 (93)   !-----
  35.0 (57)   !---
  40.0 (48)   !--
  45.0 (34)   !--
  50.0 (31)   !-
  55.0 (19)   !-
  60.0 (18)   !-
  65.0 (13)   !-
  70.0 (12)   !-
  75.0 (10)   !-
  80.0 (7)    !-
  85.0 (10)   !-
  90.0 (10)   !-
  95.0 (7)    !-
 100.0 (5)    !-
 105.0 (7)    !-
 110.0 (5)    !-
 115.0 (3)    !-
 120.0 (0)    !
 125.0 (5)    !-
 130.0 (3)    !-
 135.0 (5)    !-
 140.0 (2)    !-
 145.0 (2)    !-
 150.0 (2)    !-
 155.0 (5)    !-
 160.0 (2)    !-
 165.0 (2)    !-
 170.0 (1)    !-
 175.0 (1)    !-
 180.0 (3)    !-
 185.0 (2)    !-
 190.0 (3)    !-
 195.0 (3)    !-
 200.0 (3)    !-
 205.0 (4)    !-
 210.0 (0)    !
 215.0 (2)    !-
 220.0 (2)    !-
 225.0 (2)    !-
 230.0 (3)    !-
 235.0 (2)    !-
 240.0 (0)    !
 245.0 (2)    !-
 FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
 GROUP FREQUENCY RANGE      5.0
  TOTAL NUMBER OF OCCURRENCES=  2994 (IN RANGE)
  TOTAL NUMBER OF OCCURRENCES=  3026 (ABSOLUTE)
 0 OCCURRENCE(S) BELOW THE MINIMUM,  32 OCCURRENCE(S) ABOVE THE MA
XIMUM
 MEAN=       21.1   STANDARD DEVIATION=      37.8
 MEDIAN=     10.1   90%=      40.7   95%=     65.7
```

Part A (Continued)

```
                 STIMULUS CHARACTER COUNT
  0 (944)   !-----------------------------------------------
  4 (726)   !------------------------------------
  8 (463)   !-----------------------
 12 (299)   !---------------
 16 (259)   !-------------
 20 (140)   !-------
 24 (61)    !---
 28 (48)    !--
 32 (29)    !-
 36 (30)    !-
 40 (12)    !-
 44 (11)    !-
 48 (15)    !-
 52 (10)    !-
 56 (9)     !-
 60 (16)    !-
 64 (7)     !-
 68 (4)     !-
 72 (4)     !-
 76 (0)     !
 80 (1)     !-
 84 (1)     !-
 88 (0)     !
 92 (0)     !
 96 (0)     !
100 (0)     !
104 (1)     !-
108 (1)     !-
112 (0)     !
116 (0)     !
120 (0)     !
124 (1)     !-
128 (0)     !
132 (2)     !-
FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
GROUP FREQUENCY RANGE       4.0
 TOTAL NUMBER OF OCCURRENCES=  3094
MEAN=       11.6   STANDARD DEVIATION=       12.1
MEDIAN=      8.6   90%=       23.8   95%=       33.4
```

Part B Histograms for Length-Based Measures

```
                    ACKNOWLEDGEMENT CHARACTER COUNT
   0 (499)   !------------
   1 (1906)  !--------------------------------------------------
   2 (325)   !--------
   3 (63)    !-
   4 (10)    !-
   5 (24)    !-
   6 (16)    !-
   7 (26)    !-
   8 (4)     !-
   9 (48)    !-
  10 (1)     !-
  11 (7)     !-
  12 (3)     !-
  13 (1)     !-
  14 (61)    !-
  15 (0)     !
  16 (0)     !
  17 (0)     !
  18 (0)     !
  19 (0)     !
  20 (0)     !
  21 (0)     !
  22 (20)    !-
FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
GROUP FREQUENCY RANGE       1.0
 TOTAL NUMBER OF OCCURRENCES=  3014
MEAN=         3.2   STANDARD DEVIATION=        2.9
MEDIAN=       3.0   90%=        4.4   95%=     9.3
```

Part B (Continued)

```
                         RESPONSE CHARACTER COUNT
   0 (296)   !----------------------------------------------
   8 (277)   !-------------------------------------------
  16 (171)   !--------------------------
  24 (315)   !-------------------------------------------------
  32 (299)   !----------------------------------------------
  40 (232)   !------------------------------------
  48 (123)   !-------------------
  56 (86)    !-------------
  64 (91)    !--------------
  72 (95)    !--------------
  80 (16)    !--
  88 (22)    !---
  96 (7)     !-
 104 (14)    !--
 112 (14)    !--
 120 (13)    !--
 128 (13)    !--
 136 (6)     !-
 144 (3)     !-
 152 (7)     !-
 160 (5)     !-
 168 (2)     !-
 176 (6)     !-
 184 (1)     !-
 192 (3)     !-
 200 (4)     !-
 208 (6)     !-
 216 (5)     !-
 224 (7)     !-
 232 (4)     !-
 240 (3)     !-
 248 (3)     !-
 256 (6)     !-
 264 (1)     !-
 272 (27)    !----
 280 (24)    !---
 288 (3)     !-
 296 (7)     !-
 304 (5)     !-
 312 (7)     !-
 320 (4)     !-
 328 (8)     !-
 336 (5)     !-
 344 (2)     !-
 352 (10)    !-
 360 (6)     !-
 368 (2)     !-
 376 (8)     !-
 384 (11)    !-
 392 (5)     !-
FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
GROUP FREQUENCY RANGE      8.0
 TOTAL NUMBER OF OCCURRENCES=  2290 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  2402 (ABSOLUTE)
0 OCCURRENCE(S) BELOW THE MINIMUM,  112 OCCURRENCE(S) ABOVE THE
MAXIMUM
MEAN=       56.9   STANDARD DEVIATION=       77.3
MEDIAN=     34.1   90%=     123.0   95%=    279.8
```

Part B (Continued)

```
                STIMULUS TRANSMISSION RATE (IN CHARS PER SECOND)
  0.3 (721)  !--------------------------------
  1.8 (1120) !--------------------------------------------------
  3.3 (587)  !--------------------------
  4.8 (191)  !--------
  6.3 (72)   !---
  7.8 (45)   !-
  9.3 (18)   !-
 10.8 (8)    !-
 12.3 (2)    !-
 13.8 (1)    !-
 15.3 (1)    !-
 16.8 (1)    !-
 18.3 (0)    !
 19.8 (1)    !-
 21.3 (0)    !
 22.8 (0)    !
 24.3 (1)    !-
 25.8 (1)    !-
 27.3 (0)    !
 28.8 (290)  !------------
 FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
 GROUP FREQUENCY RANGE       1.5
  TOTAL NUMBER OF OCCURRENCES=  3060 (IN RANGE)
  TOTAL NUMBER OF OCCURRENCES=  3094 (ABSOLUTE)
 34 OCCURRENCE(S) BELOW THE MINIMUM,  0 OCCURRENCE(S) ABOVE THE MA
XIMUM
 MEAN=        5.5   STANDARD DEVIATION=       8.0
 MEDIAN=      3.6   90%=       9.5   95%=     29.1
```

Part C Histograms for Measured Rates

```
         ACKNOWLEDGEMENT TRANSMISSION RATE (IN CHARS PER SEC)
 0.3 (261)  !--------------
 1.8 (216)  !------------
 3.3 (162)  !---------
 4.8 (144)  !--------
 6.3 (124)  !------
 7.8 (122)  !------
 9.3 (63)   !---
10.8 (38)   !--
12.3 (42)   !--
13.8 (129)  !-------
15.3 (1)    !-
16.8 (15)   !-
18.3 (786)  !--------------------------------------------
19.8 (0)    !
21.3 (1)    !-
22.8 (4)    !-
24.3 (5)    !-
25.8 (2)    !-
27.3 (4)    !-
28.8 (871)  !-------------------------------------------------
FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
GROUP FREQUENCY RANGE       1.5
 TOTAL NUMBER OF OCCURRENCES=  2990 (IN RANGE)
 TOTAL NUMBER OF OCCURRENCES=  3014 (ABSOLUTE)
24 OCCURRENCE(S) BELOW THE MINIMUM,  0 OCCURRENCE(S) ABOVE THE MA
XIMUM
MEAN=       16.4   STANDARD DEVIATION=       10.4
MEDIAN=     19.3   90%=       30.5   95%=       30.7
```

Part C (Continued)

```
                    RESPONSE TRANSMISSION RATE (IN CHARS PER SECOND)
  0.3 (7)      !-
  1.8 (4)      !-
  3.3 (4)      !-
  4.8 (4)      !-
  6.3 (6)      !-
  7.8 (2)      !-
  9.3 (4)      !-
 10.8 (38)     !--
 12.3 (19)     !-
 13.8 (18)     !-
 15.3 (28)     !-
 16.8 (118)    !------
 18.3 (130)    !-------
 19.8 (86)     !----
 21.3 (201)    !-----------
 22.8 (231)    !-------------
 24.3 (325)    !------------------
 25.8 (266)    !---------------
 27.3 (37)     !--
 28.8 (859)    !--------------------------------------------------
 FREQUENCY CLASS BASE VALUE   (NUMBER OF OCCURRENCES)
 GROUP FREQUENCY RANGE       1.5
  TOTAL NUMBER OF OCCURRENCES=  2387 (IN RANGE)
  TOTAL NUMBER OF OCCURRENCES=  2402 (ABSOLUTE)
 15 OCCURRENCE(S) BELOW THE MINIMUM,  0 OCCURRENCE(S) ABOVE THE MA
XIMUM
 MEAN=       24.8   STANDARD DEVIATION=       5.0
 MEDIAN=     26.3   90%=      30.6   95%=      30.8
                                Part C  (Continued)
```